PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/31, 15/62, 1/21, C07K 14/245, C12Q 1/00	**A1**

(11) International Publication Number: WO 95/20657

(43) International Publication Date: 3 August 1995 (03.08.95)

(21) International Application Number: PCT/DK95/00042

(22) International Filing Date: 27 January 1995 (27.01.95)

(30) Priority Data:
08/187,166   27 January 1994 (27.01.94)   US

**(60) Parent Application or Grant**
  (63) Related by Continuation
  US   08/187,166 (CIP)
  Filed on   27 January 1994 (27.01.94)

(71) Applicant *(for all designated States except US)*: GX BIOSYSTEMS A/S [DK/DK]; Mothsvej 70, DK-2840 Holte (DK).

(71)(72) Applicants and Inventors: SOKURENKO, Evgeni Veniaminovic [RU/US]; Apartment 301, 1960 N. Parkway, Memphis, TN 38112 (US). HASTY, David Long [US/US]; Apartment 302, 684 Harbor Edge Circle, Memphis, TN 38103 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: KLEMM, Per [DK/DK]; Lyngbyvej 32C, DK-2100 Copenhagen Ø (DK). PALLESEN, Lars [DK/DK]; Skt. Pedersvej 4, DK-2900 Hellerup (DK). MOLIN, Søren [DK/DK]; Mothsvej 70, DK-2840 Holte (DK).

(74) Agent: PLOUGMANN & VINGTOFT A/S; Sankt Annæ Plads 11, P.O. Box 3007, DK-1201 Copenhagen K (DK).

(81) Designated States: AM, AT (Utility model), AU, BB, BG, BR, BY, CA, CN, CZ, CZ (Utility model), DE (Utility model), DK (Utility model), EE, ES (Utility model), FI, FI (Utility model), GE, HU, JP, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MX, NO, NZ, PL, RO, RU, SI, SK, SK (Utility model), TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: RECEPTOR SPECIFIC BACTERIAL ADHESINS AND THEIR USE

CLINICAL ISOLATES
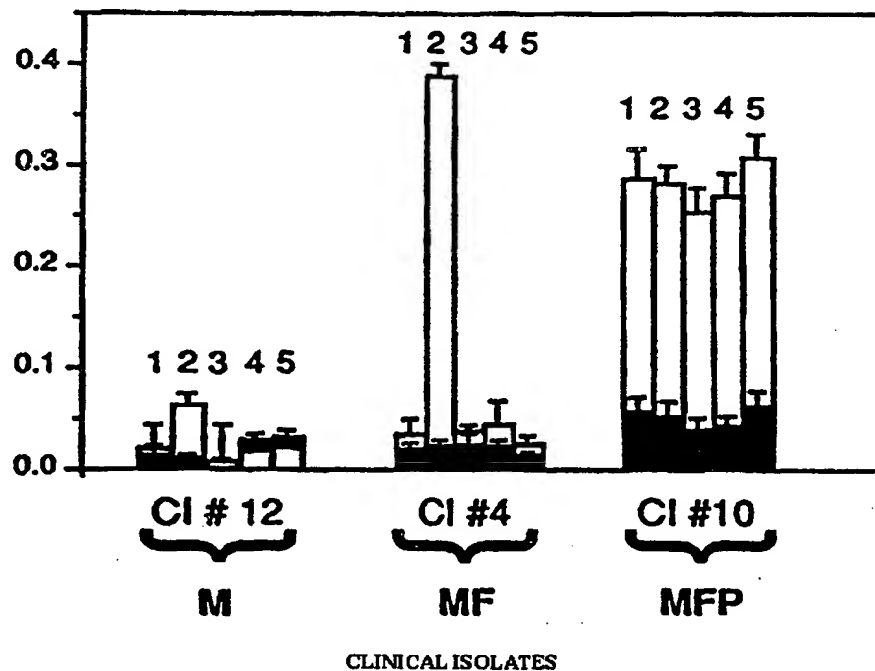
(57) Abstract

Bacterial adhesins that have been selected or recombined to have the ability to bind specifically to pre-determined, selected inanimate or animate receptors and the use of such adhesins or bacteria expressing the adhesins, in the targeting of useful compounds and/or bacteria to selected cells and surfaces.

RECEPTOR SPECIFIC BACTERIAL ADHESINS AND THEIR USE

FIELD OF INVENTION

The present invention pertains to naturally occurring bacterial adhesins and derivatives and variants hereof, having the
5 ability to bind to pre-determined, specifically selected receptors, and to the use of such adhesins in the targeting of active compounds and microbial cells to locations comprising such selected receptors.

TECHNICAL BACKGROUND AND PRIOR ART

The ability to adhere or bind specifically to, and in many
15 instances, to colonize an animate or inanimate surface is of paramount importance in microbial ecology and pathogenesis. Such specific receptor binding is provided by microbial adhesins which play a key role in bacterial/host and viral/host recognition and interaction and for the recogni-
20 tion of any specific surface by a microorganism.

Accordingly, adhesion of bacteria to host surfaces is commonly regarded as an essential step enabling bacteria to become established as members of the normal flora of host organisms or to cause an infection (refs. 7, 18). Bacterial lectins are
25 the most common and most thoroughly studied type of adhesins among both gram negative and gram positive bacteria (ref. 40). Evolutionary pressures have selected lectins for adhesive functions probably due to the abundance of glycoconjugates on animate and inanimate surfaces. One class of structures
30 that a large range of gram-positive and gram-positive bacteria including *Escherichia coli* and other members of the

family *Enterobacteriaceae*, have evolved to adhere to host glycoproteins in a saccharide-dependent manner are surface fibrils called fimbriae (ref. 14) or pili (ref. 10). Coloniz- ation Factor Antigen (CFA) type I and Colonization Factor

5   Antigen (CFA) type II are specific examples of such fimbriae.

By far the most common of the enterobacterial fimbriae is type 1, or mannose-specific (MS) fimbriae (refs. 11, 13, 14, 23). Type 1 fimbriae are heteropolymers of four different

10   subunits (refs. 28, 44). For each fimbria, about 1000 copies of a 17-kDa primary structural subunit designated FimA (or PilA), are polymerized into a right-handed helix surrounding a hollow axial core (ref. 11). Three ancillary subunits, FimF, FimG and FimH, are also polymerized into the fimbrial

15   structure, but comprise only 1-3% of the fimbrial mass (refs. 20, 24, 27, 32).

The 28 kDa FimH subunit has been shown by several direct and indirect tests to be the actual fimbrial lectin (refs. 2, 4, 20, 21, 27, 29, 32, 36, 55), although its function may be

20   affected by other subunits (ref. 55). The FimA subunit is highly variable, but the FimH subunit is highly conserved antigenically and genetically among enterobacteria (ref. 1). Interactions between type 1 fimbriae and D-mannose-containing receptors have been shown in a number of studies to play a

25   key role in the infectious process (refs. 2, 4, 9, 19, 25, 26, 31, 33, 44, 50).

Detailed analysis of adhesion-inhibition or agglutination- inhibition by various mannosides and manno-oligosaccharides have suggested that the combining site of the type 1 adhesin

30   is in the form of an extended pocket corresponding to the size of a trisaccharide and fitting best the structure α-D- Man$p$-(1-3)-β-D-Man$p$-(1-4)-D-GlcNac (ref. 16). A hydrophobic region within or close to the combining site was also pre- dicted in these studies. A similar pattern of specificity was

35   found independently in indirect adhesion-inhibition studies, as well as in direct adhesion studies using "neoglycolipids"

as receptors (refs. 37, 47). The combining site of the *Kleb-siella pneumoniae* type 1 adhesin was shown to be similar to the *Escherichia coli* adhesin, whereas the *Salmonella typhimu-rium* type 1 adhesin combining site appears to be smaller and

5   devoid of a hydrophobic region (ref. 16) Thus, it has long been thought that type 1 fimbriae of enterobacteria were functionally quite similar and that the primary essential characteristic of any potential receptor was the presence of terminal α1-3-linked mannosyl residues.

10   Recently it has been reported that the type 1 fimbriated, K-12-derived *E. coli* strain CSH-50 exhibits mannose-sensitive peptide-binding activity (ref. 51). CSH-50 *E. coli* bound to yeast mannan (Mn), a highly mannosylated glycoprotein, and to human plasma fibronectin (Fn) when immobilized on assay

15   wells. Adhesion to Mn, but not to Fn, was essentially elimin-ated by periodate treatment. Furthermore, CSH-50 *E. coli* adhered in a mannose-sensitive fashion to non-glycosylated peptide fragments of Fn and to a synthetic peptide copying the first 30 residues of the Fn molecule, FnSp1. Fimbriae

20   purified from these organisms also bound to Fn and FnSp1. A well-characterized recombinant strain of *E. coli* PC31 expres-sing type 1 fimbriae, HB101(pPKL4), adhered to Mn, but did not adhere to the other substrata. Fimbriae purified from HB101(pPKL4) did not adhere to Fn or FnSp1. Thus, *E. coli*

25   type 1 fimbriae appeared to be functionally heterogeneous.

Several *E. coli* isolates obtained from human urine also expressed peptide-binding activity similar to that of CSH-50, indicating that this new phenotype was not restricted to a laboratory strain. Other isolates expressed an adhesive

30   activity similar to that of HB101(pPKL4). A third class of type 1 fimbriae-mediated adhesive phenotype was also observed among these isolates.

The FimH subunit is the D-mannose-sensitive adhesin of type 1 fimbriae, common i.a. to the *Enterobacteriaceae*. It is pres-

35   ently widely accepted that host receptors are strictly

limited to glycoproteins containing terminal mannosyl resi-
dues (refs. 16, 37, 41, 42, 43, 47). Hereinbelow functional
and genetic evidence is provided demonstrating that this
generalization is not correct. Allelic variants of *E. coli*
5   *fimH* genes encoding proteins differing by as little as a
single amino acid substitution confer distinct adhesive
phenotypes and accordingly, the *fimH* gene is not a single
gene but rather a family of *fimH* genes.

Surprisingly, active receptors for FimH proteins were found
10 to include glycoprotein domains where mannosyl residues are
not terminal and even protein domains devoid of saccharide.
This unexpected adhesive diversity within the *fimH* family
broadens the scope of potential receptors for bacterial
adhesion and may lead to a fundamental change in the under-
15 standing of the role(s) type 1 fimbriae and other bacterial
adhesins may play in bacterial ecology or pathogenesis.

The present findings also opens up a completely new field of
technology, since it provides the means to design bacteria
expressing adhesins that bind to pre-determined, specific
20 receptors in a wide range of animate and inanimate locations.
This new technology is referred to herein as Designer Adhesin
Technology.

SUMMARY OF THE INVENTION

Accordingly, the present invention relates in one aspect to a
25 recombinant or mutant bacterial adhesin variant derived from
a naturally occurring adhesin, said adhesin variant having
altered binding properties relative to the naturally occur-
ring adhesin from which it is derived.

In further aspects the invention provides a FimH adhesin
30 having an amino acid sequence which differs from the *E. coli*
PC31 FimH adhesin by at least one amino acid and a recombi-
nant replicon comprising a DNA sequence selected from the

group consisting of a sequence coding for a recombinant
bacterial adhesin variant as defined above and a sequence
coding for a FimH adhesin as also defined above.

5   In a still further aspect, there is provided a fusion protein
comprising an adhesin selected from the group consisting of a
recombinant bacterial adhesin variant as defined above and a
FimH adhesin as also defined above, and a heterologous
polypeptide.

10   The invention also pertains to a recombinant bacterial cell
which expresses an adhesin selected from the group consisting
of a recombinant bacterial adhesin variant as defined above
and a FimH adhesin as defined above, and to a composition
comprising a population of such cells.

15   In one interesting aspect of the invention there is provided
a method of isolating a bacterial cell expressing an adhesin
having modified binding properties relative to a natively
expressed adhesin, comprising identifying in the bacterial
cell DNA sequence(s) coding for the binding domain(s) of said
natively expressed adhesin and substituting at least one
20   codon herein, whereby a modified adhesin molecule is
expressed that is different in at least one amino acid from
the adhesin expressed natively, and selecting a bacterial
cell expressing the modified adhesin having an altered
adhesion phenotype relative the natively expressed bacterial
25   adhesin.

In a further interesting aspect the invention relates to a
method of preparing a recombinant bacterial cell that binds
to a specific receptor moiety, comprising introducing into a
bacterium that does not produce an adhesin binding to said
30   receptor moiety, a DNA sequence coding for an adhesin binding
to the receptor moiety, and selecting a bacterial cell ex-
pressing the DNA sequence.

There is also provided a method of targeting a bacterial
adhesin to a specific location, comprising (i) identifying in
said location an adhesin-interacting receptor moiety which is
recognizable by bacterial adhesins, said moiety preferably
5   being one which is occurring abundantly, (ii) isolating a
bacterial cell that grows in said location and expresses an
adhesin recognizing and interacting with said receptor moie-
ty, and administering to the location the bacterial cell or
the adhesin under conditions where the adhesin and the
10   receptor moiety are brought into interacting contact whereby
the adhesin is associated with the receptor moiety.

DETAILED DISCLOSURE OF THE INVENTION

As used herein the term "bacterial adhesins" denotes proteins
which recognize and bind to a large variety of target mole-
15   cules such as polysaccharides, glycolipids, glycoproteins,
polypeptides and proteins. More than a hundred different
adhesins have been described so far originating from a large
variety of gram-negative and gram-positive bacteria. Adhesins
can be present on the bacterial surface as components of
20   organelles such as fimbriae, also called pili or fibrillae,
these three terms being used interchangeably herein, or as
non-fimbrial or afimbrial adhesins (ref. 64). Examples of
fimbrial or pili adhesins include the following surface
structures in *E. coli*: P pili, type 1 fimbriae, S pili, K88
25   pili, K99 pili, CS3 pili, F17 pili and CS31 A; in *Klebsiella
pneumoniae*: type 3 pili; in *Bordetella pertussis*: type 2
pili; in *Yersinia enterocolitica*: Myf fibrillae; in *Yersinia
pestis*: pH6 antigen and F1 envelope antigen.

Examples of non-fimbrial cell surface structures which have
30   adhesin function or which may comprise proteins having such a
function include capsules, lipopolysaccharide layers, outer
membrane proteins, NFA (non-fimbrial adhesin)-1, NFA-2, NFA-
3, NFA-4, AFA (afimbrial adhesins)-I, AFA-II and AFA-III.

In the present context, the term "fimbriae" designates long
thread-like bacterial surface organelles. Fimbriae are hete-
ropolymers each consisting of about 1000 structural compo-
nents, mostly of a single protein species. However, in many
5   cases a few percent minor components are also present. Adhes-
ins can either be identical to the major structural protein
as in *Escherichia coli* K88 and CFA1 fimbriae and type 4
fimbriae of *Pseudomonas*, *Vibrio* and *Neisseria*, or they may be
present as minor components as in *E. coli* type 1 and P
10 fimbriae [for reviews see Krogfelt 1991 (ref. 62); Kaufman
and Taylor, 1994 (ref. 60): Kuehn et al., 1994 (ref. 63);
Klemm and Krogfelt, 1994 (ref. 61)]. In the latter case, i.e.
when present as minor compounds, the adhesins are closely
related in amino acid sequence to the major fimbrial compo-
15 nent. As used herein the term bacterial adhesin will also
include adhesins isolated from non-bacterial sources includ-
ing viruses, and which are expressed in a bacterium.

In the following, the FimH adhesin of type 1 fimbriae will be
described structurally and functionally as a representative
20 example of a bacterial adhesin.

FimH is located at the tip of the type 1 fimbriae and also
intercalated at intervals in the fimbrial organelle. Most
forms of the FimH adhesin target to (bind to) oligosaccharide
structures containing terminally located α-D-mannoside resi-
25 dues [Krogfelt et al., 1990 (ref. 29)]. Based on studies with
various D-mannose derivatives the receptor binding site of
the FimH adhesin is assumed to be shaped like an elongated
pocket large enough to accommodate a trisaccharide motif
[Sharon, 1987 (ref. 65)].

30 The *fimH* gene encodes the precursor FimH protein of 300 amino
acids [Klemm and Christiansen, 1987 (ref. 27)]. This precur-
sor is processed into a mature form of 279 amino acids. The
amino acid sequence of the *E. coli* PC31 FimH protein is shown
in Table 1 below wherein cysteine residues are indicated by
35 asterixes, the signal peptide is outlined in bold letters,

and two regions contributing to the binding site are under-
lined (SEQ ID NO:1). (It should be noted that residue 176 is
a proline residue and not as previously indicated when the
PC31 FimH protein was first published, an arginine residue):

5

Table 1. Amino acid sequence of the _E. coli_ PC31 FimH protein

-21   1 *

MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVVNVG<u>QNLVVDLS</u>
         *

10 <u>TQIFCHNDYPETITDYVTLQRGSAYGGVLSNFSGTVKYS</u>GSSYPFPTTSETPRVVYNSRT

    DKPWPVALYLTPVSSAG<u>GVAI</u>KAGSLIAVLILRQTNNYNSDDFQFVWNIYANNDVVVPTG
         *   *

    GCDVSARDVTVTLPDYPGSVPIPLTVYCAKSQNLGYYLSGTHADAGNSIFTNTASFSPAQ
                                                              279
15 GVGVQLTRNGTIIPANNTVSLGAVGTSAVSLGLTANYARTGGQVTAGNVQSIIGVTFVYQ

    The FimH contains 4 cysteine residues assumed to direct
    folding of the molecule into distinct functional domains. For
    comparison FimA and the minor components FimF and FimG only
    have two cysteine residues. The localization of the cysteine
20 residues in FimH points to a tandem arrangement of two ances-
    tral genes. Furthermore, similar amino acids can be found in
    similar positions in the two halves of the FimH protein. The
    "midway" point is located roughly around residue 150 in the
    mature protein. The two halves or domains of FimH have
25 evolved differently with the N-terminal section becoming the
    domain harbouring the receptor binding site, whereas the C-
    terminal sector became the domain of the molecule required
    for integration into the fimbrial organelle structure, i.e.
    having the features of a structural component.

30 In-frame linker insertions into the _fimH_ gene confirms this
    model of the FimH protein. Thus insertions in the C-terminal
    half of the molecule generally do not interfere with the

receptor-binding ability whereas abolishment of receptor
binding ability following linker insertion in the N-terminal
is the rule (Klemm et al., unpublished data). A similar
domain structure has been observed in the PapG adhesin of P-
5   fimbriae [Hultgren et al., 1989 (ref. 59); Kuehn et al., 1994
(ref. 63)].

In accordance with the invention, the recombinant bacterial
adhesin as defined above is one which is derived from an
adhesin having certain binding properties, but which recombi-
10 nant bacterial adhesin has altered binding properties rela-
tive to the naturally occurring adhesin (the parent adhesin)
from which it is derived. As used herein this feature
encompasses situations where the adhesin variant recognizes
and binds to receptor moieties not being recognized by the
15 parent adhesin irrespective of whether the adhesin variant
has lost its normal ability to recognize and bind to a cer-
tain receptor moiety or certain receptor moieties, or not.

As used herein the term "binding" indicates that the adhesin
has a degree of affinity to the receptor moiety which enables
20 it, when brought into contact herewith, to interact in a
binding manner with this moiety whereby an adhesin-receptor
moiety association occurs. The strength of this binding
depends on the type of binding force which causes the inter-
action between the receptor moiety and the adhesin. In the
25 present context, such binding forces include covalent binding
and binding by non-covalent binding forces including hydrogen
bonds, hydrophobic interactions, van der Waal forces and
ionic interactions. Accordingly, the term "receptor moiety"
as used herein encompasses any moiety to which an adhesin may
30 interact by the above binding forces.

In one specific embodiment, the adhesin variant is a FimH
mannose-sensitive adhesin normally binding to a receptor
selected from a domain where mannosyl residues are not ter-
minal and a domain devoid of saccharide and having an amino
35 acid sequence which differs from the E. coli PC31 FimH adhes-

in by at least one amino acid residue substitution, including
an amino acid sequence differing by at least 2 amino acids,
preferably by at least 3 amino acids, more preferably by at
least 4 amino acids, most preferably by at least 5 amino

5   acids. In further useful embodiments, the amino acid sequence
may even differ by more than 5 amino acids such as at least
6, preferably by at least 7, more preferably by at least 8,
even more preferably by at least 9 and in particular by at
least 10 amino acid residues, such as by at least 12 amino

10 acids including by at least 15.

Accordingly, the above FimH adhesin variant is preferably at
least 90% homologous to the PC31 FimH adhesin, such as at
least 92% homologous, more preferably at least 93%
homologous, even more preferably at least 94% homologous,

15 most preferably at least 95% homologous, and in particular at
least 96% homologous, e.g. at least 97% homologous. In par-
ticularly interesting embodiments, the adhesin is at least
98% homologous, including at least 99% homologous such as at
least 99.5% homologous.

20 The above FimH adhesin variant can be a chimeric adhesin
comprising amino acid sequences from different FimH adhesins
having identical or different binding specificities.

As it has been mentioned above, the present invention is
generally aimed at providing the means to design bacterial

25 adhesins having specific binding properties whereby bacteria
expressing the adhesin variants or the adhesin variants in
isolated or purified form can be designed to bind to a speci-
fic desired target receptor moiety. Accordingly, the adhesin
variant may in accordance with the invention be an adhesin

30 variant as defined above which binds to an animate receptor
moiety. Such receptors include receptors located on inner
surfaces of humans and animals, such as e.g. the mucosal
membranes of the gastrointestinal tract including the teeth
and the oral cavity, and the mucosal membranes of the respir-

35 atory and the genitourinary systems. Included are also

adhesin variants that bind to outer surfaces, including the skin, of humans and animals.

In a further embodiment, the adhesin variant is designed so as to acquire the ability to bind to a plant receptor moiety.
5   This aspect is of particular interest in relation to deliber-ate release to out-door or in-door environments where plants are cultivated, of useful recombinant bacteria having a desirable effect on the growth and yield of the plants. Such desirable bacteria are e.g. bacteria expressing a pesti-
10   cidally active substance, i.e. a biopesticide including as examples pesticidal toxins produced naturally by *Bacillus* spp such as the *Bacillus thuringiensis* (Bt) toxin. In this con-text, another example is bacteria which protect plants against low temperature damages or bacteria which express
15   gene products protecting plants against detrimental effects of herbicides.

By providing such bacteria with genes expressing adhesin variants which e.g. bind specifically to certain plant spe-cies and/or to certain locations on the plant, these useful
20   bacteria will, when administered to the plant growing envi-ronment, be selectively associated with the target plant species or a specific target area of the plant. It may thus be desirable to have these useful bacteria administered to the leaves of the plants or to have the root system colonized
25   herewith.

Accordingly, the present invention encompasses adhesin vari-ants as defined herein which bind selectively or specifically to a phylloplane receptor moiety or which bind to receptors on plant roots. Similarly, adhesin variants can be provided
30   which are targeted to the stem or the flowers of the plants.

As it is mentioned above, bacterial adhesins include adhesins having an inherent capability to bind or interact with inani-mate surfaces carrying receptor moieties with which the adhesin can interact to become bound to the surfaces. It is

known that certain bacterial adhesins can bind to inanimate
surfaces including as examples glass, hydroxyapatite (a tooth
enamel model compound) or polymer structures including plas-
tics and polysilicates. The present invention has made it
5   possible to design bacteria which bind selectively to any
inanimate surface which carries a receptor moiety for which
an adhesin variant binding thereto may be constructed. Accor-
dingly, the present invention also provides an adhesin vari-
ant as defined herein which binds to an inanimate receptor
10 moiety. Such adhesin variants are particularly interesting in
connection with the concept of bioremediation, i.e. a tech-
nology designed to enhance degradation of chemical pollutants
in the environment. It is clearly a significant improvement
of this technology to have at hand bacteria which comprise
15 genes coding for pollutant-degrading gene products and which
also express adhesins targeting the bacteria selectively to
the environment where the pollutants are present, e.g soil,
aquatic environments and drinking water supply systems.
Furthermore, adhesin variants capable of binding to tooth
20 enamel are useful in the protection of teeth against caries.

In a further embodiment, there is in accordance with the
invention provided an adhesin variant which is part of a
fusion protein comprising the adhesin variant and a non-
adhesin, heterologous polypeptide. Using the FimH as an
25 example, it has been found that fusions between a bacterial
adhesin and other proteins can be made whereby the resulting
fusion proteins are inserted into the cell surface organelle
of which the adhesin is a structural part. These resulting
hybrid adhesin-carrying cell organelles remain fully func-
30 tional with respect to binding properties. Additionally, it
has been found that large regions of non-adhesin proteins,
e.g. regions comprising in the range of 1 to 100 amino acids
including a range of 5 to 75 amino acids and a range of 10 to
60 amino acids, such as regions comprising 15 to 54 amino
35 acids, can be inserted into type 1 fimbriae without impairing
the binding properties of the fimbriae.

In useful embodiments of the invention, the non-adhesin region of a fusion protein comprising an adhesin variant as defined herein include a heterologous polypeptide which is selected from an epitope, an enzyme, a toxic gene product and

5   an antibody.

It has significantly been found that, when fusion proteins are expressed in which the heterologous polypeptide is an epitope or an epitope-carrying domain forming an integrated part of the fusion protein, and thus presented on the surface

10   of the host cell expressing the fusion protein, the epitope-carrying polypeptide can be presented in a conformational form similar to its natural conformation.

Furthermore, it has surprisingly been found that the above fusion proteins can be overproduced by the bacteria compris-

15   ing hybrid genes coding for fusion proteins, resulting in excretion of the fusion proteins to the growth medium in large quantities. Accordingly, the excreted fusion proteins are then readily isolated and purified, e.g. by means of affinity chromatography. These findings provide the means to

20   manufacture bacterial cells having on their surface hybrid adhesin-carrying cell organelles as well as to produce large quantities of excreted fusion proteins, both of which can be targeted to specific surfaces as determined by the binding properties of the adhesin variant of the fusion protein.

25   The above technology of making adhesin variant-fusion proteins is useful for a range of industrially important purposes such as:

(i) development of live vaccines targeted to specific cellular surfaces;

30   (ii) development of subunit vaccines for administration orally or by injection, which are targeted to pre-determined, specifically selected cell surfaces or mucosal surfaces;

(iii) development of fusion proteins combining specific
binding properties with specific enzymatic or toxin activ-
ities. Such fusion proteins have applications as therapeuti-
cal or diagnostical agents, including use in biosensors;

5   (iv) use of fusion proteins as carriers of non-covalently
linked chemical moieties whereby the adhesin part of the
protein is used to target the chemical moiety to specific
locations and the non-adhesin part carries and then releases
the moiety when the fusion protein has reached its target.
10   Examples of chemical entities which may be linked to the
fusion protein include imaging agents and pharmacologically
active components. Examples of applications for this use
include imaging of atherosclerotic plaques or tumor tissues,
and delivery of chemical agents at specific locations in or
15   on microbial, human, animal or plant cells including specific
tissues or tissue components;

(v) development of fusion proteins which are useful in affin-
ity purification processes.

It has been found that the *fimH* gene coding for the *E. coli*
20   FimH adhesin is not a single gene but rather a family of *fimH*
genes, and accordingly it has now been established that
allelic variants of *E. coli fimH* genes exist that encode
adhesin proteins which, relative to the known *E. coli* PC31
fimH gene product differ by as little as a single amino acid
25   substitution and confer distinct binding or adhesive
phenotypes.

Accordingly, as it has been mentioned above, the present
invention relates in a further aspect to a FimH adhesin
having an amino acid sequence which differs from the *E. coli*
30   PC31 FimH adhesin as defined above by substitution of at
least one amino acid. It will be understood that such an
adhesin encompasses naturally occurring adhesins as well as
adhesins which are encoded by recombinant or mutant *fimH*
genes. In this context the term "*fimH* gene" denotes a gene

coding for a gene product which can be integrated into a type 1 fimbria and which confers to the fimbria the ability to recognize and bind to a receptor.

The FimH adhesin as defined above may be an adhesin having its inherent binding properties or an adhesin variant which in relation to an adhesin encoded by a naturally occurring gene from which the gene coding for the adhesin variant is derived, has altered binding properties. Furthermore, the FimH adhesin may be either mannose-sensitive or mannose-insensitive. The term "mannose-sensitive" is used herein to designate that the binding of an adhesin is inhibited in the presence of mannose residues. In one specific embodiment, the FimH adhesin may be a FimH adhesin normally binding to a receptor moiety selected from a domain where mannosyl residues are not terminal and a domain devoid of saccharide such as e.g. a glycolipid, a glycoprotein, a protein, a polypeptide and a peptide, including a hormone. Examples of proteins to which a FimH adhesin according to the present invention may bind include as examples animal proteins such as a casein including κ-casein, a gelatine, a globin, an albumen and a collagen, and vegetable proteins including soy protein.

The FimH adhesin according to the invention include an adhesin having an amino acid sequence which differs from the *E. coli* PC31 FimH adhesin by at least 2 amino acid residues, such as an amino acid sequence differing by at least 3 amino acids, preferably by at least 4 amino acids, more preferably by at least 5 amino acids, most preferably by at least 6 amino acids. In further useful embodiments, the amino acid sequence may even differ by more than 6 amino acids such as at least 7, preferably by at least 8, more preferably by at least 9, even more preferably by at least 10 and in particular by at least 11 amino acid residues, such as by at least 12 amino acids including by at least 15.

Accordingly, the above FimH adhesin is preferably at least
90% homologous to the PC31 FimH adhesin, such as at least 92%
homologous, more preferably at least 93% homologous, even
more preferably at least 94% homologous, most preferably at
5   least 95% homologous, and in particular at least 96%
homologous, e.g. at least 97% homologous. In particularly
interesting embodiments, the adhesin is at least 98%
homologous, including at least 99% homologous or at least
99.5% homologous.

10   In one specific embodiment, the FimH adhesin as defined above
is one which, when tested for binding to yeast mannan (Mn),
human plasma fibronectin (Fn), periodate treated Fn and the
synthetic peptide FnSp1 comprising the first 30 amino acids
of Fn, only binds to Mn. In the following, an adhesin having
15   this pattern of binding properties is designated an
M class FimH adhesin. In other specific embodiments, the FimH
adhesin is an adhesin which, when tested for binding to the
above compounds, binds to Mn and Fn (MF class FimH adhesin)
or an adhesin which among these compounds bind to all of
20   these (MFP class FimH adhesin).

It has been found that bacteria expressing FimH adhesins of
the above MFP class bind in a mannose-sensitive (MS) manner
to polyoxyethylene sorbitan monolaurate (Tween 20) and a
little less well to polyoxyethylene sorbitan monooleate
25   (Tween 80). Furthermore, bacteria expressing MFP class FimH
adhesins make a much tougher pellicle than bacteria expres-
sing other types of adhesins. In the present context, the
term "pellicle" indicates a layer or film of bacteria that
forms at the air/liquid interface of a liquid growth medium.
30   This noticeable phenomenon might be of particular interest
where there is a reason to concentrate microorganisms at the
surface of an aquatic environment, such as e.g. bacterial
cells which in accordance with the present invention express
a pollutant-degrading gene product.

Another interesting finding is that bacteria expressing a MFP class adhesins bind to hydroxyapatite to a higher degree than do bacteria expressing a M class adhesin. Hydroxyapatite, especially saliva-treated hydroxyapatite is i.a. used as a
5   model for tooth enamel, and accordingly, this finding indicates that bacteria expressing MFP class adhesins are particularly useful in bacterial compositions intended for colonization of teeth.

It has also been found that the MFP class adhesins bind to a
10   large range of synthetic peptides and accordingly seem to have a broad specificity in terms of amino acid motifs.

In further specific embodiments of the invention, the FimH adhesin is an adhesin which, when tested for binding to the five Fn-fragments obtained by thermolysin treatment as it is
15   described in reference No. 51, only binds to the 40-kDa gelatin-binding fragment or which binds to all of these Fn-fragments, or to none of these.

In addition to the above classes of FimH adhesins, another class has been identified which is designated the M^L (low
20   adhesive) class. Such an adhesin confers the ability to aggregate yeast cells in a mannose-sensitive (MS) fashion, in titers similar to M class adhesins, but surprisingly, it binds at only low levels to Mn or Fn and FnSp1. Furthermore, adhesins of this low adhesive M^L class adhere poorly to MDCK,
25   buccal cells and erythrocytes as compared with M class adhesins. Example of a M^L class adhesin is one expressed by the recombinant *E. coli* strain KB 23 which differs only from the PC31 FimH adhesin by having an alanine instead of a valine at residue 27 and the FimH adhesin expressed by the human fecal
30   *E. coli* isolate which is designated F-18 [McCormick et al., 1989 (ref. 34)]. This latter adhesin differs from the PC31 FimH in three amino acid residues and the F-18 isolate has been found to colonize the large intestine to a higher degree than certain *E. coli* K-12 strains do. Accordingly, it is
35   contemplated that these M^L class adhesins confer to

gastrointestinal bacteria the ability to colonize the large
intestine which is significant for a live bacterial vaccine
for exerting its immunological effect in the gastrointestinal
tract.

5 Furthermore, it has been found that among M class adhesins
adhesion is found that is not sensitive to inhibition by D-
mannose. Such a mannose-insensitive (or mannose-resistant) M
class adhesin is designated in the following as an M^R adhes-
in. One example of a bacterial strain expressing an M^R adhes-
10 in is the clinical isolate U221-3 which is mentioned in the
following.

In accordance with the invention, a FimH adhesin as defined
above can be a chimeric adhesin comprising amino acid
sequences from different FimH adhesins. Such chimeras are
15 constructed e.g. by providing multiple restriction fragments
of a *fimH* gene, followed by exchanging under ligation condi-
tions these fragments with corresponding fragments of an
other *fimH* gene and cloning the ligation product as it is
described in Example 1 below. As it is also explained below,
20 recombinant plasmids containing such chimeric *fimH* genes can
be transformed into a host cell and transformants tested for
adhesive phenotype, allowing determination of the regions of
each gene capable of conferring functional activity (Fig. 5).
These studies which are described in details below showed
25 that all of the sequence changes relative to the PC31 *fimH*
gene that affected binding function in the studied strains of
E. coli CSH-50 and clinical isolates (CIs) designated #s 3,
4, 7, 10, F-18 and U221-3, respectively, occurred between
residues 27 and 119, both included, of the 279 residue,
30 mature *fimH* sequences.

Accordingly, the invention encompasses in one embodiment a
FimH adhesin comprising an amino acid sequence which differs
from the *E. coli* PC31 FimH adhesin by at least one amino acid
occurring between residues 27 and 119 of the mature FimH
35 sequence, including a FimH adhesin comprising an amino acid

sequence which differs from the *E. coli* PC31 FimH adhesin by
at least one amino acid occurring between residues 33 and 78
of the mature FimH sequence.

The selected potential receptors for a FimH adhesin as
5   defined above include those animate and inanimate receptors
mentioned above for a recombinant bacterial adhesin variant
and the potential uses of the FimH adhesins are also the same
as those uses described above for this recombinant bacterial
adhesin variant.

10   As mentioned above, the invention relates in a further aspect
to a recombinant replicon comprising a DNA sequence coding
for a recombinant bacterial adhesin variant as defined herein
or a DNA sequence coding for a FimH adhesin as also defined
herein. Such a replicon is suitably selected from a chromo-
15   some or a plasmid. The DNA sequence includes a sequence which
is inserted by conventional recombination techniques such as
insertion by means of restriction enzymes and subsequent
ligation, or the DNA sequence is provided by subjecting a
replicon comprising a naturally occurring sequence coding for
20   an adhesin to a mutagenization procedure including site-
directed mutagenesis, insertion of a transposable element,
mutagenization by radiation or chemical mutagenization,
followed by selection of cells comprising a mutated sequence
conferring altered binding properties relative to a cell
25   comprising the wild-type sequence.

In preferred embodiments, the recombinant replicon is one
having a broad host range including bacterial species na-
turally occurring in soil, in aquatic environments, on inner
and outer surfaces of humans and animals, and which is com-
30   patible with replicons occurring in potential host strains.

In one useful embodiment, the recombinant replicon as defined
above is one wherein the DNA sequence codes for a FimH adhes-
in having an amino acid sequence which differs from the *E.
coli* PC31 FimH adhesin by at least one amino acid, including

an adhesin having an amino acid sequence which differs from the *E. coli* PC31 FimH adhesin by at least 2 amino acid residues, such as an amino acid sequence differing by at least 3 amino acids, preferably by at least 4 amino acids, more

5 preferably by at least 5 amino acids, most preferably by at least 6 amino acids. In further useful embodiments, the amino acid sequence may even differ by more than 6 amino acids such at least 7, preferably by at least 8, more preferably by at least 9, even more preferably by at least 10 and in particu-

10 lar by at least 11 amino acid residues, such as by at least 12 amino acids including by at least 15.

Accordingly, the above recombinant replicon preferably comprises a DNA sequence coding for a FimH adhesin which is at least 90% homologous to the PC31 *fimH* gene, such as at least

15 92% homologous, more preferably at least 93% homologous, even more preferably at least 94% homologous, most preferably at least 95% homologous, and in particular at least 96% homologous, e.g. at least 97% homologous. In particularly interesting embodiments, the adhesin is at least 98%

20 homologous, including at least 99% homologous such as at least 99.5% homologous.

In a further embodiment, the above replicon comprises a DNA sequence which is a chimeric *fimH* gene as it has been defined above, comprising DNA from different *fimH* genes. The replicon

25 can also be one which comprises a further DNA sequence e.g. derived from a microorganism selected from a bacterium, a virus, a protozoan, a fungus and a yeast. This further DNA sequence is e.g. one coding for a heterologous polypeptide, including an epitope, an antibody, a toxic gene product, an

30 enzyme, a pesticidally active gene product and a pollutant-degrading gene product.

In useful embodiments, the replicon as defined herein comprises a DNA sequence which is isolated from an *Enterobacteriaceae* species, including a DNA sequence which is isolated

from *E. coli*, a *Klebsiella* sp., an *Enterobacter* sp., a
Yersinia sp. or a *Salmonella* sp.

In addition to being a DNA sequence as defined above, the
sequence can be a synthetic sequence constructed by conven-
5   tional techniques of DNA synthesis.

As it is also mentioned above, the present invention
encompasses a fusion protein comprising a recombinant bacte-
rial adhesin variant or a FimH adhesin as defined above, and
a heterologous polypeptide. Such a polypeptide is in useful
10   embodiments an immunologically active gene product i.e. an
epitope (antigenic determinant) from a pathogenic organism,
which polypeptide, when administered to the body of a human
or an animal is capable of stimulating the formation of
antibodies therein. A cell in which such an epitope is
15   expressed is advantageously utilized in the preparation of
live vaccines. Such vaccines have several advantages over
known live vaccines:

Firstly, the epitope forms a structural part of an adhesin
which is embedded in a surface organelle of the vaccine
20   cells. This implies that the hybrid DNA sequence coding for
the epitope further comprises the means for transporting the
epitope, when expressed, to the outer surface of the cell,
i.e. translocating it through the cell membrane. This is
immunologically highly advantageous, since the epitope will
25   be brought more closely in contact with immunologically
competent cells of the body to which the fusion protein-
expressing vaccine cells are administered.

Secondly, the adhesin part of the epitope-carrying fusion
protein can be selected so as to have specific binding prop-
30   erties whereby the vaccine cell may be targeted to a particu-
lar location in the body where an immunological response to
the epitope is desirable. The adhesion of the epitope-carry-
ing cell to a particular location or region of the body will
in this manner ensure that the cell is retained in the human

or animal body in that particular location for a period of
time which is sufficient to obtain the desired immune
response.

In accordance with the invention, a useful cell for expres-
5   sion of the above fusion protein is one selected from a
bacterial species which inherently contains an adhesin-carry-
ing surface organelle. Such species include as examples gram-
negative species of *Enterobacteriaceae* such as *E. coli*,
Klebsiella spp, *Salmonella* spp, *Yersinia* spp, *Vibrionaceae*,
10 *Hemophilus* spp, *Bordetella* spp and *Pseudomonadaceae*, and
gram-positive species such as *Neisseria* spp and *Streptococcus*
spp.

The epitope part of a fusion protein according to the inven-
tion can be an epitope derived from any pathogenic organism
15 or agent against which it is desirable to develop vaccines.
Such pathogenic organisms include viruses, bacteria and
eucaryotic organisms such as fungi, yeast or protozoa.

Whereas cells expressing an epitope-carrying fusion protein
as defined herein may be used as a live vaccine, it is also
20 within the scope of the invention to provide isolated and/or
purified cell surface organelles comprising the fusion pro-
tein, including fimbriae and pili, as a vaccine, and it is
also contemplated that useful vaccines may be provided where-
in cells expressing an epitope-carrying fusion protein have
25 been killed by conventional methods such as formaldehyde
treatment or thermal treatment.

In a further embodiment of the invention, the fusion protein
according to the invention comprises as the non-adhesin
polypeptide part a toxic gene product e.g. having a selective
30 toxic effect on particular cells in the body such as e.g.
cancer cells. By selecting the adhesin part as one having a
specific binding affinity to receptors in such cells it is
possible to have cells expressing the toxic gene product
bound selectively to such target cells whereby these cells

may be killed or damaged by the toxic gene product. It is also possible to use isolated or purified cell organelles containing a fusion protein comprising the cell toxic (cytotoxic) gene product for the purpose of targeting the
5 toxic product.

In a further interesting embodiment, the fusion protein comprises an antibody. Such an embodiment is, *inter alia*, particularly interesting with respect to the provision of fusion proteins which may be used in affinity purification of
10 biological compounds having binding affinity to the antibody part of the fusion protein. It is contemplated that cells expressing as part of a surface organelle, such a fusion protein may be utilized directly as a means of concentrating a biological compound, or the isolated surface organelles
15 comprising the antibody-carrying fusion protein may be used for this purpose.

Furthermore, the fusion proteins as defined herein are useful as carriers of non-covalently bound compounds such as pharmacologically active, diagnostically active and imaging com-
20 pounds with the purpose of providing cells or cell organelles carrying the active compounds, which thereby become targetable to particular regions or locations of a body to which these cells or cell organelles are administered. The invention encompasses any combination of a fusion protein as
25 defined herein and an active compound which can be covalently bound to a fusion protein.

As mentioned above, the present invention encompasses in one aspect a recombinant bacterial cell which expresses a recombinant bacterial adhesin variant or a FimH adhesin as defined
30 above. In one specific embodiment, the bacterial cell is one which comprises the above-defined recombinant replicon. Depending on the field of application of such a cell, it may e.g. be selected from a soil bacterium, an aquatic bacterium, a bacterium which is normally associated with plants, a
35 bacterium which is member of the human or animal indigenous

bacterial flora, or a bacterium which is adapted to colonize certain ecological niches such as e.g. sewage purification plants or certain inanimate surfaces.

The major significant advantages which have been achieved by the present invention is the possibility to provide recombinant bacterial cells which are not only ecologically well-adapted to grow in a particular ecological environment, but which are also provided with means for colonizing more permanently in their ecologically natural environment. These means for improved ability to colonize an environment are the adhesins expressed by the bacteria which have been constructed and/or selected so as to enable the recombinant bacterial cell to adhere to or bind to specific receptors in the environment, i.e. the bacterial cells are targeted to that environment. Thereby the bacteria according to the present invention will have an ecologically competitive advantage relative to organisms in the particular environment which do not have surface structures comprising adhesins binding to receptors present in the environment, at least not to the same extent as the bacterial cells according to the invention.

In addition to the environment-specific adhesins which the bacterial cell expresses, the cell will have a phenotype which is desirable in the environment to which it is targeted. As one example, a cell according to the invention which is originally isolated from the human or animal indigenous bacterial flora may typically be one which expresses an epitope including an epitope which is part of a fusion protein expressed by the bacterial cell. As another example may be mentioned a bacterial cell which is isolated from a plant and which expresses a pesticidally active compound such as a *Bacillus thuringiensis* toxin. Further examples include a plant root-associated nitrogen-fixating bacterium isolated from soil which in accordance with the invention is provided with adhesins improving the capability of the bacterium to become permanently colonized to the roots of a specific plant

or specific plants, or a bacterium which is ecologically associated with an aquatic or terrestrial environment containing pollutants to be degraded or removed.

Accordingly, the recombinant bacterial cell can be derived from any gram-negative or gram-positive bacterium for which a need exists to obtain improved colonization in a particular inanimate or animate environment. Such bacteria include as examples *Enterobacteriaceae* spp, *Hemophilus* spp, *Neisseria* spp, *Bordetella* spp, *Streptococcus* spp, *Pseudomonadaceae* spp, *Vibrionaceae* spp, *Baccilaceae* spp.

In certain embodiments of the invention it is advantageous that the present recombinant bacterial cell is provided as one which, when it is administered to a particular location or environment, will not persist in that environment. Accordingly, such a recombinant bacterial cell may further comprise a gene coding for a gene product which, when expressed has a killing or cell function-limiting effect in said cell, the expression of said gene coding for the cell killing or cell function-limiting gene product being regulated in such a manner that the bacterial cell when targeted to receptor in a specific location will be killed or limited in its function in a pre-determined manner. The gene coding for the cell killing or cell function-limiting gene product is suitably regulated by a factor selected from the group consisting of a stochastic event, the presence/absence of a chemical compound in the location, and a physical factor.

In a further aspect, the invention relates to a method of isolating or constructing a recombinant bacterial cell expressing an adhesin having modified binding properties relative to a natively expressed adhesin such as a natively expressed FimH adhesin. As it is defined above, this method comprises identifying in the bacterial cell DNA sequence(s) coding for the binding domain(s) of said natively expressed adhesin and substituting at least one codon herein whereby a modified adhesin molecule is expressed that is different in

at least one amino acid from the adhesin expressed natively, and selecting a bacterial cell expressing the modified adhesin having an altered adhesion phenotype relative to the natively expressed bacterial adhesin.

5   As it is explained in details below, the binding domain can e.g. be identified by constructing chimeric adhesin-encoding genes and screening for cells which by having a region in the adhesin gene replaced by a corresponding heterologous region of a different DNA sequence, acquires a new binding
10   phenotype. Having identified a binding domain of the natively expressed adhesin, recombinant cells having desirable binding phenotypes may be obtained by substituting one or more codons in the binding domain(s) to obtain expression of recombinant adhesins and selecting cells having the desirable phenotypes.
15   The substitution of codons may be achieved by methods know per se such as site-directed mutagenesis using synthetic oligonucleotides and PCR technology or transposable elements or by conventional radiation or chemical mutagenization.

      In certain useful embodiments, the above method includes
20   steps whereby a non-adhesin compound is associated with the adhesin, e.g. a step where a gene coding for the recombinant adhesin is part of a hybrid gene comprising a gene coding for a non-adhesin polypeptide which thereby is expressed with the recombinant adhesin as part of a fusion protein comprising
25   the adhesin. Furthermore, recombinant adhesins resulting from the above method may in specific embodiments comprise a non-covalently bound compound which is associated with the adhesin when expressed.

      As mentioned above, the invention also encompasses recombi-
30   nant bacterial cells having selected binding properties whereby cells with desirable phenotypes can colonize environments where the presence of bacteria having a particular phenotype is advantageous. Accordingly, there is in a further aspect of the invention provided a method of preparing a
35   recombinant bacterial cell that binds to a specific receptor

moiety, comprising introducing into a bacterium that does not
produce an adhesin binding to said receptor moiety, a DNA
sequence coding for an adhesin binding to the receptor moie-
ty, and selecting a bacterial cell expressing the DNA

5   sequence.

The primary objective of this method is to provide the means
of constructing a bacterial strain having the capacity to
colonize an environment, based on a parent strain which has
an inherent, useful phenotype in this particular environment

10   but which does not express an adhesin binding to receptor
moieties in the environment. Accordingly, the method includes
as a first step the isolation of an environmentally adapted
bacterium not binding to appropriate receptor moieties and in
subsequent steps, the identification of heterologous genes

15   encoding adhesins which bind to receptor moieties occurring
in said environment, preferably moieties occurring abundant-
ly, isolating this gene and introducing it into the above
parent strain. The adhesin gene may e.g. be a gene coding for
a naturally occurring FimH adhesin or a recombinant FimH

20   adhesin as defined above.

In one useful embodiment of the method, the adhesin-encoding
gene is introduced by transforming a parent bacterial cell
with a recombinant replicon as defined herein. In further
embodiments, the method is designed so as to obtain a cell

25   wherein a non-adhesin compound is associated with the adhes-
in, e.g. by introducing the gene coding for an adhesin as a
hybrid gene coding for a non-adhesin polypeptide whereby non-
adhesin compound is expressed with the adhesin as part of a
fusion protein comprising the adhesin, or by binding non-

30   covalently a compound to the adhesin when expressed.

Besides the above method, an adhesin carrying bacterial cell
having an altered pattern of adhesion can be provided by
using a selection procedure comprising contacting an appro-
priately sized population of wild-type adhesin-carrying

35   bacterial cells with a potential receptor moiety to which the

wild-type cells do not adhere, e.g. in a manner as it is
disclosed in Example 6 below whereby spontaneously or random-
ly mutated cells having acquired the ability to adhere to the
receptor moiety in question, become progressively enriched.
5   From such an enriched culture, cells with the new adhesion
ability can readily be isolated and further characterized.

As it has been explained in details above, one primary objec-
tive of the present invention is to provide the means of
targeting a compound to a specific location. Accordingly, the
10 invention relates in an important aspect to a method of
targeting an adhesin to such a location. The method comprises
the identification in the location of a receptor moiety, said
moiety preferably being one which occurs abundantly in the
particular location, which moiety can recognize and interact
15 with an adhesin, and the isolation of a bacterial cell which
is capable of growing in the location and expressing an
adhesin which recognizes and interacts with the identified
receptor moiety, and administering the cell or the adhesin in
an isolated form to that location.

20 The identification of a suitable receptor moiety in a par-
ticular location can be carried out in several manners. One
example is a screening procedure where cells expressing known
adhesins or known isolated adhesins are administered to the
location e.g. being isolated cells or tissues of microbial,
25 animal or plant origin or an inanimate surface as defined
herein, and screening for binding/adhesion of the tested
adhesins e.g. according to adhesion assays as disclosed
herein. If binding of one or more adhesins occurs, it is an
indication that receptor moieties for that or those tested
30 adhesin(s), is/are present in the location.

Alternatively, available data with regard to the presence and
amounts of chemical moieties present on the surfaces of the
location may be collected or such data have to be generated,
and based upon such data, adhesins which are known to bind to
35 one or more of the identified major moieties are selected and

their binding to this/these structure(s) is tested e.g. according to the assays as used herein. Chemical moieties which are considered potential adhesin-interacting receptor moieties include as examples glycolipids, glycoproteins,
5   proteins, polypeptides, saccharide moieties and peptides.

In the case no suitable chemical moiety is identified in the location, which is capable of binding to known adhesins or which bind with a sufficient affinity, it is required to construct a library of modified adhesin molecules based on
10 known adhesins which are modified by replacing one or more codons as it is explained herein, and/or such a library provided by constructing synthetic adhesin molecules, and then screening this library for recognition of and interaction with identified location surface moieties. A library of
15 modified FimH adhesins may e.g. be selected for specificity towards a given receptor by running clones of these adhesins through a column or matrix containing the receptor moiety in question or cells or tissues isolated from the location without knowing what the receptor moiety is. The clone(s)
20 expressing the adhesins with affinity to receptor moiety/moieties will adhere/bind to the column or matrix, and can subsequently be isolated therefrom.

It is within the contemplation of the invention that crystallographic analyses of adhesins, whether naturally occurring
25 or constructed as indicated above, is a useful technique for the obtainment of information about adhesin structures that assumingly will recognize and interact with particular adhesin receptor moieties.

In accordance with the invention, one embodiment of the above
30 method is one wherein the isolated bacterial cell expresses an adhesin having modified receptor moiety-binding properties relative to an adhesin natively expressed by the cell, the isolation of the cell comprising identifying in a parent bacterial cell, DNA sequence(s) coding for the binding
35 domain(s) of said natively expressed adhesin and substituting

at least one codon herein, whereby a modified adhesin mole-
cule is expressed that is different in at least one amino
acid from the adhesin expressed natively, and selecting a
bacterial cell expressing the modified adhesin having an
5   altered adhesion phenotype relative to the natively expressed
bacterial adhesin or a method wherein the bacterial cell
expressing an adhesin that recognizes and binds to the
receptor moiety is a recombinant bacterial cell derived from
a parent bacterial cell that does not produce an adhesin
10 binding to said receptor, by inserting into the parent cell a
DNA sequence coding for an adhesin binding to the receptor
moiety, and selecting a bacterial cell expressing the DNA
sequence.

One primary objective of the present invention is the target-
15 ing of useful non-adhesin compounds to a particular location.
Accordingly, the invention encompasses in an interesting
embodiment a method as defined above wherein a non-adhesin
compound is associated with the adhesin, whereby said com-
pound is targeted with the adhesin to the location comprising
20 the receptor moieties recognizable by the adhesin.
The compound can be associated with the adhesin by a covalent
binding or by any of the above mentioned non-covalent types
of molecule interaction forces.

When associated covalently with the adhesin the compound to
25 be co-targeted to the selected location with the adhesin can
be an enzyme, an antibody, an epitope or a toxin which is
part of a fusion protein comprising the adhesin. A compound
which is associated with the adhesin by a non-covalent bind-
ing is typically a pharmacologically active, diagnostically
30 active or imaging compound.

Locations to which it is desirable to have an adhesin tar-
geted by the present method include a human or animal sur-
face, a plant surface and an inanimate surface as defined
above.

In one specific embodiment of the present method the bacteri-
al cell being administered to the location expresses a recom-
binant bacterial adhesin variant derived from a naturally
occurring parent adhesin, said adhesin variant having altered
5 binding properties relative to the naturally occurring adhes-
in from which it is derived, the altered binding properties
including binding to at least one receptor moiety to which
the parent adhesin does not bind. Such an adhesin variant is
advantageously derived from a naturally occurring adhesin
10 isolated from a cell structure selected from the group con-
sisting of a capsule, a lipopolysaccharide layer, on outer
membrane protein, a flagellum, a pilus, a fimbria, a non-
fimbrial adhesin (NFA) or an afimbrial adhesin (AFA).

In specific embodiments of the invention, the above adhesin
15 variant as used in the present method is a protein having an
amino acid sequence differing in at least one amino acid
residue from its parent protein adhesin such as a FimH adhes-
in having an amino acid sequence which differs from the *E.
coli* PC31 FimH adhesin as defined herein in at least one
20 amino acid. Such a FimH adhesin includes an adhesin which
binds to a receptor selected from the group consisting of a
domain where mannosyl residues are not terminal and a domain
devoid of saccharide and an adhesin variant which is at least
90% homologous to the PC31 FimH adhesin as defined herein,
25 such as at least 92% homologous, more preferably at least 93%
homologous, even more preferably at least 94% homologous,
most preferably at least 95% homologous, and in particular at
least 96% homologous, e.g. at least 97% homologous. In par-
ticularly interesting embodiments, the adhesin is at least
30 98% homologous, including at least 99% homologous or at least
99.5% homologous.

The above FimH adhesin can be a chimeric adhesin as defined
above, comprising amino acid sequences from different FimH
adhesins and constructed according to the methods below.

In accordance with the invention, an adhesin can be adminis-
tered to a location in the form of an adhesin-expressing
bacterial cell. Such a cell is one capable of growing in that
particular location. Accordingly, the bacterial cell is

5 suitably derived from a bacterial species which is normally
occurring in the location including human or animal body
surfaces, plant surfaces such as plant root surfaces and
inanimate surfaces. In this context, an animal body surface
includes the insect gut, whereto it is desirable to adminis-

10 ter a bacterial cell expressing an insecticidally active
toxin.

Thus, if it is desired to administer the bacterial cell to
the root of a plant, a suitable bacterial cell is preferably
isolated from a strain which has colonized the rhizosphere of

15 that plant to a large degree, i.e. the strain is a major
member of the natural plant root flora. Such an isolate is
then provided with a gene coding for an adhesin which will
recognize and interact with an abundantly occurring moiety on
the roots of said plant. In this manner, a suitable adhesin

20 which is expressed naturally in a bacterium which is not
adapted to grow in a plant rhizosphere, becomes expressible
in a normal inhabitant of the rhizosphere environment (loca-
tion).

In specific embodiments of the present method of targeting a

25 bacterial adhesin to a specific location, the adhesin is a
FimH adhesin as defined above, having an amino acid sequence
which differs from the E. coli PC31 FimH adhesin as defined
herein in at least one amino acid.

In an interesting embodiment, the adhesin-carrying bacterial

30 cell being targeted is a cell which further comprises a gene
coding for a gene product which, when it is expressed, has a
killing or cell function-limiting effect in said cell, the
expression of said gene coding for the cell killing or cell
function-limiting gene product being regulated in such a

35 manner that the bacterial cell, when targeted, will be killed

or limited in its function in a pre-determined manner. The expression of such a "suicide" or cell function-limiting gene may suitably be regulated by a factor selected from the group consisting of a stochastic event, the presence/absence of a

5   chemical compound in the location and a physical factor. As examples of such "suicide" or cell function-limiting genes providing the means of biological containment, may be mentioned those disclosed in WO 87/5932 and WO 93/20211

Furthermore, the present Designer Adhesin Technology (DAT)
10 provides very useful means of obtaining colonization with desirable bacteria in a particular environment with the purpose of obtaining beneficial changes of the microbial flora in the environment. As one example, certain bacterial species in the gastrointestinal (GI) tract of humans and
15 animals have beneficial effects on the health condition of the host organism e.g. by suppressing pathogenic organisms or by contributing to the digesting of certain diet components. The present technology makes it possible to select particularly useful bacteria from the GI-tract and have them
20 designed in accordance with the present invention, to have improved colonization abilities. Similar examples include desirable bacterial colonizations of biological sewage purification systems, plants where invasion of pathogenic organisms may be controlled by colonizing the plants with harmless
25 bacteria, and teeth where caries may be controlled by colonizing the dental enamel with bacteria suppressing those causing the caries attacks.

In another industrially interesting aspect, the invention provides the means of isolating a compound from a solution or
30 suspension containing the compound. The method comprising contacting the solution or the suspension with a fusion protein as defined herein wherein the heterologous polypeptide has an affinity to the compound to be isolated.

Furthermore, the invention provides a composition comprising
35 a population of a bacterial cell as defined herein.

The invention is further illustrated in the below Examples
and the Figures, wherein

Fig. 1 is a schematic model for the construction of recom-
binant plasmids pGB1-24 (containing *fimH* from CI #10) and
5   pGB2-24 (containing *fimH* from PC31) used for transforming *E.
coli* AAEC191A(pPKL114) with cloned *fimH* genes. Plasmid pGB2-
24 was used as the vector for all other cloned *fimH* genes
described herein;

Fig. 2 is a restriction map of *fimH* genes. Five unique res-
10   triction sites are present in the PC31 *fimH* gene. Numbers in
parentheses following enzymes are the base pair positions of
the cut sites. Some of these sites are found in the other
fimH genes, as marked. Chimeric genes were produced by ex-
changing each available restriction fragment from the other
15   five *fimH* genes with corresponding fragments in the PC31 gene
and then recombinant strains expressing resulting chimeric
fimH subunits were tested for adhesion. Fragments indicated
by boxes are those which conferred MF or MFP adhesive
phenotypes on the chimeric genes;

20   Fig. 3 illustrates adhesion of representative "wild-type" (A)
and recombinant (B) M-class, MF-class and MFP-class strains
to Mn (1), Fn (2), periodate-treated Fn (3) and to FnSp1 (4).
Strain designations given for the "wild-type" strains are
given in AS. Strain designations KB31, KB12, KB4, KB7, KB50
25   and KB10, are for recombinant strains of AAEC191A(pPKL114),
which is *fimH⁻*, after transformation with plasmids that
contain *fimH⁺* from strains HB101(pPKL4), CI #12, CI #4, CI
#7, CSH-50 and CI #10, respectively. Open columns indicate
results when bacteria were incubated in buffer without D-
30   mannose, while solid columns are results in the presence of
D-mannose. Values indicated are the mean ± S.E.M. (n=4) for
each column;

Fig. 4 illustrates the adhesion of representative M-class,
MF-class and MFP-class strains (CIs #12, #4 and #10, respect-

ively) to Fn fragments prepared by thermolysin treatment as
described in ref. 51. Columns labelled 1-5 indicate adhesion
to: 1) NH2-terminal 30-kDa domain; 2) the 55-kDa gelatin-
binding domain; 3) the 110-kDa cell attachment domain; 4) the
5   29-38-kDa heparin binding domains; and 5) the 20-kDa COOH-
terminal domain. Open columns represent adhesion in the
absence of D-mannose; solid columns represent adhesion in the
presence of D-mannose. Mean ± S.E.M. (n=4);

Fig. 5 is a composite figure illustrating comparison of amino
10   acid sequences of FimH adhesins and active restriction frag-
ments of *fimH* genes. The published nucleotide and deduced
amino acid sequence of the PC31 *fimH* gene and gene product
(ref. 27) serve as prototype. Numbered amino acid residues
shown above the model of the PC31 FimH represent residues
15   that are different in other FimH subunits due to amino acid
substitution or deletion. Standard one-letter code applies
and residues in the other FimH sequences that are different
are indicated. Deleted amino acids are indicated by Δ. It
should be noted that residue 176 is not arginine as published
20   previously (ref. 27) for the PC31 FimH, but proline. Regions
of the FimH subunits conferring change in adhesive phenotype,
highlighted in bold, were determined by functional assays
performed on chimeras between the "classic" mannose-specific
PC31 *fimH* gene present in HB101(pPKL4) and the above
25   described genes. Residues predicted to be key in conferring
receptor specificity are circled. Approximate positions of
unique restriction sites used to create chimeras are indi-
cated along the bottom of the model;

Fig. 6 illustrates plasmid pPKL4 which is a derivative of
30   pBR322 (thick line) carrying the entire *fim* operon (FimA-H)
including the regulatory genes *fimB* and *fimE* (not shown), and
the promoter region with the *SnaBI* site. In this plasmid an
8mer linker with an *BglII* site was inserted in the *SnaBI* site
to create pPKL83;

Fig. 7 illustrates the construction of plasmid pSM1314; the vector pVLT33 is a derivative of the broad host range replic-on RSF1010. Plasmid pPKL83 was digested with *Bgl*II and pVLT was digested with *Bam*H1; the two were ligated and pSM1314 was

5 the resulting plasmid in which expression of the fimA-H cluster is under the control of the *tac* promoter;

Fig. 8 illustrates plasmid pLPA22 and derivatives hereof as used in this study. The triangles indicate the position of translational stop-linkers in the *fimH* gene in plasmid

10 pPKL115. The positions of heterologous inserts are indicated (black boxes). Small triangles indicate signal-peptide encoding sectors.

Fig. 9 illustrates plasmids pLPA29, pLPA30, pLPA36, pLPA58, pLPA59 and pLPA98;

15 Fig. 10 shows immuno-electron microscopy with colloid gold labelling of *E. coli* HB101 cells containing plasmids pLPA22 plus pPKL115 (a), pLPA37 plus pPKL115 (b), pLPA38 plus pPKL115 (c), using anti-pre-S2 monoclonal antiserum. Bar, 0.1 μm.

20 EXAMPLE 1

Functional heterogeneity of type 1 fimbrial adhesins due to minor sequence variations among *fimH* genes

1.1. Materials and methods

1.1.2. Reagents

25 Yeast Mn, a polymannosylated glycoprotein isolated from *Saccharomyces cerevisiae* cell walls, was obtained from a commercial source (Sigma Chemical Co, St. Louis, MO, U.S.A.). Mannan is composed of an N-linked backbone of β1,2-linked mannopyranose units with α-linked mannopyranose side chains

30 (ref. 38). The majority of the carbohydrate of human plasma

Fn is composed of N-glycosidic complex-type biantennary glycans and no high mannose-type or hybrid-type N-glycans have been described (refs. 30, 45, 54). Human plasma Fn and Fn fragments were purified as described previously (refs. 5,
5   15, 51, 58). Periodate treatment was performed as described previously (ref. 51). The synthetic peptide, FnSp1, copying the first 30 amino acid residues of the Fn molecule (EAQQMVQ-PQSPVAVSQSKPGCYDNGKHYQI) was synthesized in the Protein Chemistry Laboratory of the VA Medical Center, Memphis, TN
10   (SEQ ID NO:2). The saccharide content of the four substrata was characterized using two lectins, concanavalin A (ConA), well known to react with terminal and internal mannosyl residues, and the *Calanthus nivalis* agglutinin (GNA), which recognizes only terminal Manα1-3Man, Manα1-6Man and Manα1-
15   2Man sequences (E. Y. Laboratories, San Mateo, CA). Immobilized Mn and Fn both reacted with ConA, whereas GNA bound only to Mn. These results are consistent with the known structures of the oligosaccharide moieties of these two compounds. Neither lectin reacted with immobilized FnSp1.
20   Periodate treatment (ref. 51) of Mn or Fn eliminated lectin reactivity.

1.1.3. Bacterial strains and plasmids

The CSH-50 strain (lambda⁻,F⁻araΔ(lac-pro) rspL thi fimE::IS1) is a Cold Spring Harbor K12-derived strain (ref.
25   35). The *E. coli* strain MG 1655 (CGSC6300; K12 derivative, lambda⁻,F⁻) and a derivative strain AAEC191A (MG1655 recA Δfim were generously provided by Dr. Ian Blomfield (Bowman Gray University, Winston-Salem, NC). AAEC191A has had the entire *fim* gene cluster deleted by allelic exchange (ref. 8).

30   Clinical isolates (CIs) were urinary tract isolates obtained from the clinical microbiology laboratories of the Memphis VA Medical Center or The City of Memphis Hospitals, Memphis, TN. The 12 CIs used in this study were selected on the basis of MS agglutination of yeast cells after growth in broth, a
35   classic test for type 1 fimbriae.

Plasmid pPKL4, a pBR322 derivative containing the entire *fim*
gene cluster from *E. coli* strain PC31 (K12-derivative, *gal
tonA phx ara*) and encoding for the expression of fully func-
tional type 1 fimbriae in HB101 *(supE hsdS recA ara proA lacY*
5   *galK rspL xyl mtl ΔfimBE)*, has been described previously
(ref. 28). pPKL114 is a recombinant plasmid derived from
pPKL4, but with a translational stop-linker inserted into the
*Kpn*l site in the *fimH* gene. No transcriptional effects of the

stop-linker are to be expected. Antibiotics were used at the
10   following final concentrations: ampicillin (50 μg/ml),
kanamycin (60 μg/ml) and chloramphenicol (30 μg/ml).

1.1.4. Polymerase chain reaction

Oligonucleotide primers were designed using the published
sequence for the *fimH* gene in pPKL4 (ref. 27). The 5' primers
15   copied regions 13 and 49 bp upstream from the *fimH* gene and
were extended on the 5' end by an *Apa*ll restriction site and
a GC clamp: Primer 1: 5'-GGGGG-GTGCAC-ACC TAC AGC TGA ACC
CGG-3' (SEQ ID NO:3); Primer 2: 5'-GGGG GTGCAC T CAG GGA ACC
ATT CAG GCA-3' (SEQ ID NO:4). The 3' primers copied 18 bases
20   of the bottom strand of the *fimH* gene that encode for the 6
terminal amino acids of *fimH* and were extended by an *Fsp*l or
*Sph*l site and a GC clamp: Primer 3: 5'-GGG TGCGCA TTA TTG
ATA AAC AAA AGT CAC - 3' (SEQ ID NO:5); Primer 4: 5'-GGG
GCATGC TTA TTG ATA AAC AAA AGT CAC-3' (SEQ ID NO:6). Primer 1
25   and 3 were used for CI #10 and pPKL4, primer 1 and 4 were
used for CI #4 and CSH-50 and primer 2 and 4 were used for CI
#s 7 and 12 to generate PCR products from plasmid or
chromosomal DNA prepared from *E. coli* expressing different
functional classes of type 1 fimbriae. The PCR reaction
30   mixture consisted of template DNA, primer pairs, dNTPs, and
Taq DNA polymerase in PCR buffer. The PCR was performed in a
Perkin-Elmer Cetus automatic thermal cycler with denaturation
at 96°C for 1 min., primer annealing at 55°C for 1 min., and
primer extension at 72°C for 2 mins. for a total of 40
35   cycles. All of the PCR products migrated similarly in agarose

gels. Purification, restriction and ligation of DNA was
performed using standard procedures (refs. 39, 48). All
primers for PCR and for nucleotide sequencing were produced
by the Molecular Resources Center, UT, Memphis.

5 1.1.5. Subcloning

The PCR products from CI#10 and from pPKL4 were cut with
respective restriction enzymes and ligated into the *Apal*1 and
*Fsp*1 restriction sites of plasmid pACYC177 (New England
Biolabs, Beverly, MA, U.S.A.) which is compatible with the
10 pBR322-based pPKL114 to be used in complementation experi-
ments, creating plasmids pGB1 and pGB2, respectively (Fig.
1). However, it became inconvenient to use pACYC177-based
plasmids because of a high frequency of appearance of sponta-
neous Km^r in the AAEC191A host strain. The origin of this
15 problem is not entirely clear, but it was avoided by subclon-
ing the *fimH* genes from pGB1 and pGB2. The inserts and
upstream regions of pACYC177 containing the *tet* promoter were
cut from pGB1 and pGB2 with *Fsp*1 and *Bam*H1 and subcloned into
the polylinker site of pGEM-3Z (Promega, Madison, WI) that
20 had been cut with *Bam*H1 and *Hinc*2, creating plasmids pGB11
and pGB2-1 respectively. pGEM-3Z was simply used as a con-
venient intermediate in subcloning into pACYC184.

The inserts were cut out again using *Sma*1 and *Hind*3 and
subcloned into pACYC184 (New England Biolabs, Beverly, MA)
25 cut with *Hinc*2 and *Hind*3, creating plasmids pGB1-2 and pGB2-
24 containing the *fimH* genes from CI#10 and pPKL4, respect-
ively. These plasmids complement the non-adhesive defect of
AAEC191A(pPKL114) giving the adhesive phenotypes of the
parental strains (see Results). Cutting the *fimH* gene from
30 pGB2-24 using *Apal*1 and *Sph*1 makes it possible to easily
insert other *fimH* genes obtained by amplifying chromosomal
DNA of other isolates by PCR. All recombinant strains we have
tested thus far using this technique exhibit the same adhes-
ive phenotype as the parent strains from which the *fimH* genes
35 were cloned.

1.1.6. Construction of chimeric *fimH* genes

Unique restriction sites (Fig. 2) were used to construct chimeric *fimH* genes between the prototypical MS pPKL4 *fimH* gene, used as genetic background, and restriction fragments
5 obtained from the newly described *fimH* genes. Fragments were purified from agarose gels and ligated into restriction "spaces" generated in the pPKL4 *fimH* gene present in pACYC184 (pGB2-24). Each chimera was analyzed by restriction mapping and the nucleotide sequences of bridging segments were deter-
10 mined to ensure proper constructions. The plasmids containing chimeric *fimH* genes were transformed into AAEC191A(pPKL114) and clones were tested for agglutination of yeast cells and for adhesion to Mn, Fn and FnSp1.

1.1.7. Nucleotide sequencing

15 The nucleotide sequences of *fimH* genes were determined by the dideoxynucleotide chain termination method of Sanger (ref. 49) using a Sequenase II® kit (U.S. Biochemical Corp., Cleveland, Ohio) and [α-^{35}S]dATP (800 to 1,000 Ci/mmol) according to the manufacturer's suggestions. The amino acid sequences
20 were deduced from nucleotide sequences using MacVector® protein and DNA analysis software (Eastman Kodak, Rochester, NY). To ensure fidelity of the PCR amplification, selected *fimH* genes were re-amplified, cloned, tested for activity and re-sequenced. More recently, we have used the *fmol*™ Polymer-
25 ase Sequencing System (Promega, Madison, WI), because it is useful with small amounts of DNA and thus subcloning the *fimH* genes from the pACYC184-based plasmids to high copy number plasmids was obviated. Bands were visualized by autoradiography of sequencing gels and compared with the published
30 *fimH* gene sequence (ref. 27).

1.1.8. Yeast cell aggregation assay

E. coli were tested for their ability to aggregate yeast cells. Commercial baker's yeast, *Saccharomyyces cerevisiae*,

was suspended in PBS (5 mg dry weight/ml). *E. coli* were
washed in PBS, resuspended to an OD_{530} of 0.4, and mixed with
the yeast cell suspension in PBS with or without 1% D-
mannose. Aggregation was monitored visually and the titer

5   recorded as the last dilution giving a positive aggregation
reaction.

1.1.9. Adhesion assays

Adhesion assays were performed as described previously (ref.
51). Briefly, microtiter assay wells were coated with

10   receptor molecules as indicated in the text and figure leg-
ends. After the wells were washed two times with PBS, 100 μl
bacterial suspensions were added in 0.1% BSA-PBS. After
incubation at 37°C for indicated times, wells were washed
three times with PBS and adherent bacteria were detected by

15   using rabbit anti-*E. coli* serum. Antibody binding was
detected using peroxidase-conjugated goat anti-Rabbit IgG.
Reaction product generated from the 5-aminosalicylic acid
substrate was measured at 405 nm after 10-15 minutes by using
an automatic microplate reader (Molecular Devices, Inc.,

20   Menlo Park, CA). Values reported are corrected for background
reaction using BSA coated plates as control.

1.2. Results

In a previous publication it was reported that type 1
fimbriae of *E. coli* CSH-50 and HB101(pPKL4) differ func-

25   tionally in their pattern of adhesion to Mn, Fn, periodate-
treated Fn and a synthetic peptide, FnSp1, immobilized on
plastic (ref. 51). Since CSH-50 and HB101(pPKL4) are labora-
tory strains, we tested 12 clinical *E. coli* isolates (CIs)
obtained from human urine for adhesion to these four substra-

30   ta. All of the CIs agglutinated yeast cells in a MS fashion.
Five of the twelve CIs adhered only to Mn. The adhesive
activity of HB101(pPKL4) and of CI #12 are shown as examples
of this class, which we have tentatively designated as M
class (Fig. 3A). Three of the 12 CIs adhered to Mn and Fn,

but not to periodate-treated Fn or to FnSp1. The adhesive
activities of CI #s 4 and 7 are shown as examples of this
class, designated as MF class. Three of the twelve CIs
adhered to each of the substrata. The adhesive activities of
5   CSH-50 and CI #10 are shown as examples of this class, desig-
nated as MFP class.

Adhesion of strains representing these three classes to Fn
fragments further illustrates the distinct differences
between the three classes. The M class CI #12 does not adhere
10   to any of the Fn fragments (Fig. 4). The MF class CI #4
adheres to the 40-kDa gelatin-binding fragment. The MFP class
CI #10 adheres, with only slight differences, to all 5 frag-
ments of Fn tested. Periodate treatment eliminated binding of
CI #4 to domain 2, but had no effect on the binding of CI #10
15   to any of the Fn domains (data not shown).

Since the *fimH* subunit has been shown to mediate the mannose-
sensitive activity of type 1 fimbriae, we focused our initial
efforts to elucidate the molecular basis for the observed
functional heterogeneity on the fimH gene. *fimH* genes were
20   amplified from chromosomal (or plasmid, for pPKL4) DNA and
the genes were cloned into pACYC177 and subcloned into
pACYC184 under control of the β-lactamase promoter of
pACYC177, according to Materials and Methods (Fig. 1)

The adhesive phenotypes conferred by the *fimH* genes were
25   tested in the following way. *E. coli* K-12 strain AAEC191A
(Δfim) was first transformed with plasmid pPKL114, which
contains an intact *fim* gene cluster but with a translational
stop-linker inserted into the last gene, *fimH*. This deriva-
tive produces morphologically normal fimbriae that are non-
30   adhesive due to absence of the FimH subunit. Plasmids har-
bouring cloned *fimH* genes were transformed into *E. coli*
AAEC191A(pPKL114) and the resultant strains were tested for
their ability to adhere to Mn, Fn, periodate-treated Fn and
to FnSp1 (Fig. 3B). Each of the recombinant strains displayed
35   adhesive phenotypes mimicking those of the representative

parent strains from which the *fimH* genes were obtained. *fimH* genes were cloned from each of the other 8 CIs and similar results were obtained with the adhesion of recombinant strains mimicking that exhibited by the parental CIs.

5   The complete nucleotide sequences of each of the six repre-
     sentative *fimH* genes were determined and the amino acid
     sequences of the *fimH* proteins were deduced as it is shown in
     Table 1 below which is a representation of amino acid
     sequences of the FimH subunits deduced from nucleotide
10   sequences of selected *fimH* genes disclosed in this example
     [CI#3 (SEQ ID NO:33), CI#4 (SEQ ID NO:29), CI#7 (SEQ ID
     NO:30), CI#10 (SEQ ID NO:31) and CI#12 (SEQ ID NO:28)] and
     those of the *E. coli* K12 strain PC31 (SEQ ID NO:1) and *E.
     coli* strain CSH-50 (SEQ ID NO:32). Additionally, the FimH
15   amino acid sequences of the following clinical isolates of *E.
     coli* are shown: KB21 (SEQ ID NO:27), KS54 (SEQ ID NO:35),
     U221-3 (SEQ ID NO:36), MJ#9-3 (SEQ ID NO:37), MJ#31-3 (SEQ ID
     NO:38), MJ#11-2 (SEQ ID NO:39), MJ#2-2 (SEQ ID NO:1) and F-18
     (SEQ ID NO:34). Standard one-letter code applies. Deleted
20   amino acid residues are indicated by Δs. M, M^L, MF, MFP, and
     M^R indicate the adhesin class as defined above.

Table 1. Amino acid sequences of the FimH proteins deduced from nucleotide sequences of *fimH* genes of clinical isolates disclosed in this example and of *E. coli* K12 strains PC31 and CSH-50

```
         -21        -16                                    1                                              20
PC31     M K R V I T L F A V L L M G W S V N A W S F A C K T A N G T A I P I G G G S A N V M
KB21     - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
CI #12   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
CI #4    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MT
CI #7    - - - - - - N - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MT
CI #10   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MTP
CSH50    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MTP
CI #3    - - - - - - N - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
F-18     - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
KS-54    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
U221-3   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mn
MJ#9-3   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#31-3  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#11-2  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#2-2   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M


         21          27          33                                              58        61
PC31     Y V N L A P V V N V G Q N L V V D L S T Q I F C H N D Y P E T I T D Y V T L Q R G M
KB21     - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
CI #12   - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
CI #4    - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
CI #7    - - - - - - A - - Ⓗ - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MT
CI #10   - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MT
CSH50    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MTP
CI #3    - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - Ⓡ - - - MTP
F-18     - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
KS-54    - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
U221-3   - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#9-3   - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mn
MJ#31-3  - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#11-2  - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#2-2   - - - - - - A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M


         62          70     73        78                                              102
PC31     S A Y G G V L S N F S G T V K Y S G S S Y P F P T T S E T P R V V Y N S R T D K P M
KB21     - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
CI #12   - - - - - - - S - - - - - - - N - - - - - - - - - - - - - - - - - - - - - - - Mt
CI #4    - - - - - - - S - Ⓔ - - - - - N - - - - - - - - - - - - - - - - - - - - - - - M
CI #7    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - L - - - - - - - - - MT
CI #10   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MT
CSH50    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MTP
CI #3    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MTP
F-18     - - - - - - - S - - - - - - - N - - - - - - - - - - - - - - - - - - - - - - - Mt
KS-54    - - - - - - - S - - - - - - - N - - - - - - - - - - - - - - - - - - - - - - - Mt
U221-3   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#9-3   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mn
MJ#31-3  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#11-2  - - - - - - - S - - - - - - - N - - - - - - - - - - - - - - - - - - - - - - - M
MJ#2-2   - - - Ⓓ - - - S - - - - - - - N - - - - - - - - - - - - - - - - - - - - - - - M


         103                          116 - 119                                      143
PC31     W P V A L Y L T P V S S A G G V A I K A G S L I A V L I L R Q T N N Y N S D D F ..
KB21     - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
CI #12   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
CI #4    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MT
CI #7    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MT
CI #10   - - - - - - - - - - - (Δ Δ Δ Δ) - - - - - - - - - - - - - - - - - - - - - - MTP
CSH50    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MTP
CI #3    - - - - - - - - - - - - - - - - Ⓥ - - - - - - - - - - - - - - - - - - - - - MTP
F-18     - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
KS-54    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mt
U221-3   - - - Ⓡ - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#9-3   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mn
MJ#31-3  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#11-2  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ#2-2   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
```

and gene products should also be identical, subunit incorporation into the fimbrial superstructure should not vary significantly. These results emphasize that in these experiments it is the FimH subunit that determines receptor specificity.

5   In comparing the new FimH sequences to the one published previously (ref. 27), the only structural alteration that can be clearly linked to a functional change, without resorting to analysis of chimeric *fimH* genes, is the non-conservative substitution of arginine[58] in the MFP class CSH-50 FimH

10   subunit for leucine[58] in the M class PC31 FimH subunit. Since each of the other FimH sequences had more than one change, it was necessary to construct chimeric genes to begin to focus on functionally relevant changes.

In the case of the CI #10 FimH, an MFP class adhesive activ-

15   ity similar to that of CSH-50 is conferred by a different region of the gene which encodes for a subunit deleted of residues 116-119. It remains to be determined how two distinctly different structural changes can bring about apparently similar changes in receptor specificity. It is pos-

20   sible, of course, that as additional receptor molecules are tested, these two variants will be found to be functionally distinct.

The *ApaI1-Tth111I* fragment of the CI #7 *fimH* gene confers MF class activity in the CI#7/PC31 *fimH* chimera. Since the

25   asparagine[16]-threonine[16] substitution is within the leader sequence and thus not represented in the mature protein, the histidine[33]- asparagine[33] substitution must be of functional importance for the MF class CI #7 FimH. Comparison of the active regions of the MF class CI #4 and the M class CI #12

30   FimH subunits suggests the importance of the glutamic acid[73]-glycine[73] substitution for MF class activity of the CI #4 FimH. Thus, histidine[33], arginine[58], glutamic acid[37] and deleted glycine[116]-isoleucine[119] appear to be key residues in the functional activity of FimH subunits of CI #7, CSH-50, CI

35   #4 and CI #10, respectively, but a more precise demonstration

of which residues are involved and how they affect the ligand-binding cleft(s) remains to be performed.

At first glance, the FimH mediated, mannose-sensitive pro-
tein-binding activity of type 1 fimbriae is the most surpris-
5   ing of the adhesive phenotypes described here. However,
protein-binding activity of FimH (i.e. PilE) subunits was
noted earlier in a study characterizing *mutT*-induced muta-
tions in the *fimH (pilE)* gene (Harris et al., ref. 22) How-
ever, the protein-binding activity described by Harris et al.
10   was not mannose-sensitive. It is presently not known whether
the protein-binding activity described herein is in addition
to or separate from the mannose-binding activity, but the
concept of bifunctional properties of lectins has been estab-
lished for several years (ref. 6). While the MFP class type 1
15   fimbriae appears to react somewhat promiscuously with most Fn
fragments, the reaction does not appear to be non-specific.
For instance, the MFP class CSH-50 type 1 fimbriae do not
adhere well to gelatin (ref. 51). Further, the adhesion to
ovalbumin is sensitive to both periodate and glycosidase
20   treatment (ref. 51). Further work will be required to deter-
mine the consensus amino acid motif reactive with this class
of FimH subunit.

Previous studies suggested that the combining site of the *E.
coli* FimH adhesin is in the form of an extended pocket cor-
25   responding to the size of a trisaccharide with an associated
hydrophobic region (ref. 16). The MS nature of all of the
adhesive interactions described suggests that if the combin-
ing sites are separate, they may be close to each other.
However, it remains to be determined whether or not the
30   mannose effect is direct or allosteric. Conformational
changes that frequently occur in lectins upon binding the
saccharide ligand (ref. 46) could affect a second, distant
binding site. Site-directed mutations may be sufficient to
clarify which structural changes result in changes in
35   receptor specificity. However, such studies are unlikely to
shed much light on how the structural changes actually relate

Table 1, continued

```
        144                               163   166                      184
PC31    F V W N I Y A N N D V V V P T G G C D V S A R D V T V T L P D Y P G S V P I P L T M
KB21    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
CI #12  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
CI #4   - - - - - - - - - - - - - - - - - - - - - - Ⓒ - - - - - - - - - - - - - - - - - M
CI #7   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
CI #10  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MF
CSH50   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MF
CI #3   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MFP
F-18    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MFP
KS-54   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
U221-3  - - - - - - - - - - - - - - - - - - - - - Ⓗ - - - - - - - - - - - - - - - - - - ML
MJ09-3  - - - - - - - - - - - - - - - - - - - - - Ⓗ - - - - - - - - - - - - - - - - - - M
MJ031-3 - - - - - - - - - - - - - - - - - - - - Ⓗ - - - - - - - - - - - - - - - - - - - Mr
MJ011-2 - - - - - - - - - - - - - - - - - - - Ⓐ Ⓗ - - - - - - - - - - - - - - - - - - - M
MJ02-2  - - - - - - - - - - - - - - - - - - - Ⓐ - - - - - - - - - - - - - - - - - - - - M
```

```
        185                        201                                   225
PC31    V Y C A K S Q N L G Y Y L S G T H A D A G N S I F T N T A S F S P A Q G V G V Q L M
KB21    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
CI #12  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
CI #4   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
CI #7   - - - - - - - - - - - - D - - - - - - - - - - - - - - - - - - - - - - - - - - - M
CI #10  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MF
CSH50   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MF
CI #3   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MFP
F-18    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MFP
KS-54   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
U221-3  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
MJ09-3  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ031-3 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mr
MJ011-2 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ02-2  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
```

```
        226                                                             266
PC31    T R N G T I I P A N N T V S L G A V G T S A V S L G L T A N Y A R T G G Q V T A G M
KB21    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
CI #12  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
CI #4   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
CI #7   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
CI #10  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MF
CSH50   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MF
CI #3   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MFP
F-18    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - MFP
KS-54   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
U221-3  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - ML
MJ09-3  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ031-3 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Mr
MJ011-2 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
MJ02-2  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M
```

```
        267                       279
PC31    N V Q S I I G V T F V Y Q M
KB21    - - - - - - - - - - - - - ML
CI #12  - - - - - - - - - - - - - M
CI #4   - - - - - - - - - - - - - MF
CI #7   - - - - - - - - - - - - - MF
CI #10  - - - - - - - - - - - - - MFP
CSH50   - - - - - - - - - - - - - MFP
CI #3   - - - - - - - - - - - - - ML
F-18    - - - - - - - - - - - - - ML
KS-54   - - - - - - - - - - - - - M
U221-3  - - - - - - - - - - - - - Mr
MJ09-3  - - - - - - - - - - - - - M
MJ031-3 - - - - - - - - - - - - - M
MJ011-2 - - - - - - - - - - - - - M
MJ02-2  - - - - - - - - - - - - - M
```

The nucleotide and deduced amino acid sequences of the pPKL4 *fimH* gene are identical to that reported previously, except that residue 176 is not an alanine residue as previously reported, but a proline residue. Independent re-amplifica-
5 tion, re-cloning and re-sequencing confirmed this sequence for the pPKL4 *fimH* gene. Sequencing was also repeated on independently amplified and cloned isolates of the CI #10 and CI #7 fimH genes to confirm sequence fidelity and no errors were found.

10 The nucleotide and deduced amino acid sequences of the other *fimH* alleles described in this Example are > 98% conserved, but there is more than one amino acid residue difference in all but one of the new *fimH* sequences when compared to the published pPKL4 sequence. To focus on the sequence differ-
15 ences that resulted in changes in functional activity, advantage of unique restriction sites were taken (Fig. 2) to construct chimeric *fimH* genes. Multiple restriction fragments covering the entirety of each of the sequenced *fimH* genes were exchanged with corresponding fragments in the prototypi-
20 cal *fimH* gene of E. *coli* strain PC31 that was amplified from pPKL4, cloned into pACYC184 and used as the genetic background. Recombinant plasmids containing the chimeric *fimH* genes were transformed into E. *coli* AAEC191A(pPKL114) and transformants were tested for adhesive phenotype, allowing
25 determination of the regions of each gene capable of conferring functional activity (Fig. 5). All of the sequence changes that affected function occurred between residues 33 and 119 of the 279 residue mature *fimH* sequence.

1.3. Discussion

30 The functional heterogeneity which is described above must be due entirely to allelic variants of the *fimH* gene. The only variables in the recombinant strains which are described in this Example are the *fimH* genes; all other genes necessary for fimbrial subunit synthesis, transport and assembly are
35 the same in each case. Since the ratios of the various genes

to the ligand-binding cleft(s) and it will ultimately be necessary to determine the 3-dimensional structure of FimH or FimH fragments crystallized in the presence of ligand to fully understand structure/function relationships.

5   The three adhesive classes of type 1 fimbriae identified above may understate the functional heterogeneity of type 1 fimbriae. The group of CIs that has been tested in this Example is small and only a few substances have been tested as potential receptors. A larger group of isolates tested

10   against additional receptor candidates might yield additional functional classes. Preliminary studies with MS *Enterobacter aerogenes* and *Klebsiella pneumoniae* strains exhibiting MF class and MFP class activity suggest that heterogeneous receptor specificities will also be found among other type 1

15   fimbriated enterobacterial species.

It is also believed that it is possible that adhesins from some fimbriae responsible for mannose-resistant hemagglutina-tion or adhesion are structurally related to FimH, but with sequence alterations that eliminate sensitivity to mannose.

20   The possibility that the MS lectin-like properties of FimH might be eliminated while retaining other adhesive properties of FimH (e.g. pellicle formation) has been shown previously in a study characterizing *mutT*-induced mutations in the *fimH (pilE)* gene (ref. 26). At the minimum, it is believed that

25   tests for type 1 fimbriation should include additional func-tional characterization. While all type 1 fimbriae-mediated adhesion which have been described in this Example is mannose-<u>sensitive</u>, it is not all mannose- or even saccharide-<u>specific</u> as has commonly been thought. Further studies of

30   type 1 fimbriae as a virulence factor must be able to distin-guish among the various functional classes.

Allelic variation of the so-called G adhesins of P fimbriated uropathogenic *E. coli* also results in different functional classes, but the requirement for the Galα1-4Gal sequence

35   within isoreceptors is maintained (refs. 52, 53). These

differences in G adhesin receptor specificity appear to be
rather subtle, at least superficially, when compared to the
differences in FimH receptor specificities. Yet there is
significantly greater sequence homology among the *fimH* genes
5   than among the G adhesin genes, some of which share less than
50 percent homology. The G adhesin receptor specificities
affect host susceptibility, due in large part to host-speci-
fic expression of glycolipid isoreceptor variants. Whether
the FimH family of adhesins bears a similar relationship to
10   host susceptibility or tissue tropism remains to be deter-
mined. In this regard it is possible that the G adhesin
family could exhibit additional receptor specificities not
restricted to the Galα1-4Gal sequence. The lectin-independent
affinity of P fimbriae for immobilized Fn is not dependent on
15   the G adhesin, but on two other minor subunits, E and F,
neither of which bear significant homology to FimH (refs. 56,
57).

It is important to point out that the degree of functional
heterogeneity of type 1 fimbriae described in the present
20   Example was not appreciated when any of the studies cited
above were performed. The results of these studies have made
it clear that structural and functional heterogeneity occurs
within the class of adhesive organelles commonly referred to
as MS or type 1 fimbriae and that the adhesive diversity will
25   lead to a broader spectrum of receptive surfaces for poten-
tial colonization. The surprising finding that a FimH family
of adhesins exists may prove to be an important step toward
unravelling the role(s) type 1 fimbriae may play in the
ability of enterobacteria to reach their normal habitat or
30   gain entry into deeper tissues, where devastating effects can
occur.

EXAMPLE 2

Expression of type 1 fimbriae in heterologous bacterial species

The *fim* operon of *E. coli* comprises a cluster of genes
5   covering about 8 kb of DNA. This operon has been isolated and
cloned on plasmids in its entirety. The promoter upstream of
the *fimA* gene is located within an invertible DNA sequence,
which in *E. coli* leads to a switch on/switch off situation
for fimbrial synthesis. In one orientation of the invertible
10   sequence the promoter is directed towards the *fim* genes, and
the cell is fimbriated; in the other orientation the promoter
is directed in the opposite direction, and the cell is non-
fimbriated.

Since the regulation of the switch of the invertible promoter
15   sequence is very complex and involves several genes outside
the *fim* operon it is far from certain that the switching
takes place in other bacteria than the enterics. It was
therefore considered necessary to insert a replacement promo-
ter for the expression of the *fim* genes, and as a model for
20   gene expression in a number of different bacterial species
the *lac* promoter was chosen. This promoter has been shown to
be active and regulatable in many bacterial species.

Plasmid pPKL83 is a derivative of pPKL4 (ref. 27) carrying
the entire *fim* operon in pBR322, in which the promoter has
25   been destroyed by inserting a *Bgl*II linker in the *SnaB* site
located in the promoter sequence. There is a second *Bgl*II
site in plasmid pPKL83 upstream of the *fim* operon (Fig. 6).
Plasmid pVLT33 (Fig. 7) is a kanamycin resistant derivative
of the broad host range plasmid RSF1010, carrying the *lacI*^{q}
30   gene and the *tac* promoter placed upstream of a multiple
cloning site in which a unique *Bam*H1 site is placed. The two
plasmids were ligated together after digestion of pPKL83 with
*Bgl*II and pVLT33 with *Bam*H1. In one orientation (pSM1314),
this fusion plasmid will express fimbriae in the presence of

IPTG due to the fusion between the fimbrial genes and the *lac* promoter.

The correct orientation of the fusion plasmid pSM1314 was verified by transforming it into a strain of *E. coli* which
5　carries a deletion of the *fim* operon. Production of fimbriae was assayed in two ways: 1) Cell aggregation with fimbrial antibodies and 2) ELISA assay of whole cells. The former analysis is rather simple: to a small volume (10 μl) of an outgrown or IPTG-induced culture of the cells to be tested is
10　added a small volume (2 μl) of antibodies raised against fimbriae, on a glass slide. After mixing the samples, fimbriated cells begin to show cell aggregates which are easily observed directly as clumps or under a microscope. No clumping was observed with cells of the strain with a *fim*
15　deletion, whereas pSM1314 transformants of this strain showed clearly detectable cell aggregates. The ELISA analysis of whole cells confirmed the aggregation assay. In Table 2 below the readings from this type of assay are presented, and they show quantitatively the occurrence of Fim antigens on the
20　cells as a result of IPTG induction of the pSM1314 carrying strain.

Table 2. Results (duplicate) of ELISA assay for type 1 fimbria expressed by pSM1314 in E. coli AAEC191 (OD$_{492}$)

AAEC191(pSM1314)	0.145/0.164
AAEC191(pSM1314)+IPTG	1.026/1.260
Blank	0.113/0.095

25

Plasmid pSM1314 also carries a *mob* site which allows it to be transferred to other gram negative bacteria provided a helper plasmid is introduced. This type of transfer is most easily
30　performed in "triparental" matings in which a donor strain (*E. coli* carrying pSM1314), a helper strain (*E. coli* carrying a plasmid with conjugation genes) and a recipient strain carrying a selectable marker not present in any of the two

other strains, are mixed on a plate (directly or on a fil-
ter). After some growth (often overnight) this mixture is
spread on selective plates with antibiotics that only allow
the recipient carrying the desired plasmid to grow and form
5   colonies.

In the present context, the *E. coli* strains MC1000(pSM1314)
and MC1000(pRK2013) and (as recipient) *Enterobacter cloacae*
strain A50 Nal^r (ref. 67), were mated. This recipient strain
is resistant to nalidixic acid. After selection for growth on
10   plates with kanamycin plus nalidixic acid the resulting
clones were grown in liquid medium and assayed for the pres-
ence of fimbriae in the absence/presence of IPTG. The cell
aggregation assay was employed.

This assay showed that fimbriae were produced in the *Enterob-*
15   *acter cloacae* strain and were present on the cell surface;
however, full repression of expression from the *tac* promoter
was not obtained, most likely due to an increased escape
synthesis. The results showed that *E. coli* type 1 fimbriae
may be synthesized and processed correctly for pili formation
20   on the surfaces of heterologous gram-negative bacterial
species.

The plasmid pSM1314 in *E. coli* HB101 was deposited on 26
January 1994 with DSM, the Deutsche Sammlung von Mikroor-
ganismen und Zellkulturen GmbH, (German Collection of Micro-
25   organisms and Cell Cultures), Mascheroder Weg 1B, D-38124
Braunschweig, Germany, under the accession number DSM 8922.

EXAMPLE 3

<u>The construction of *fimH*-fusion genes and the expression of</u>
<u>mannose-sensitive FimH fusion proteins</u>

30   Heterologous sequences mimicing the pre-S2 region of the
hepatitis B viral surface antigen and a neutralizing epitope

of the cholera toxin B chain were inserted in two different
positions in the FimH adhesin of type 1 fimbriae. This was
carried out by introduction of restriction site handles
(BglII-sites) in the fimH gene, followed by in-frame inser-
5   tion of heterologous DNA segments encoding the foreign epito-
pes. In the selected positions such insertions did not sig-
nificantly alter the adhesive function of the FimH protein,
since hosts producing hybrid fimbriae that contained the
chimeric adhesins exhibited adhesion phenotypes and were
10   normally fimbriated. The heterologous inserts of 52 and 15
amino acids, respectively, residing in the chimeric FimH
proteins were recognized by specific sera on the surface of
the fimbriae on bacterial hosts. The results illustrate the
possibility of using bacterial adhesins as general presenters
15   of foreign antigens and epitopes.

3.1. Materials and methods

3.1.1. Bacterial strain and growth conditions

The Escherichia coli K12 strain HB101 was used in this study
as a host for expression of chimeric fimbriae. This strain is
20   phenotypically Fim⁻ due to a deletion in the fim gene cluster
(ref. 8). Cells were grown on solid medium or in liquid broth
supplemented with appropriate antibiotics. When required,
gene expression from the lac promoter, residing in front of
the fimH gene in plasmid pLPA22 and its derivatives, was
25   induced by the addition of IPTG (isopropyl thiogalacto-
pyranoside) to the growth medium.

3.1.2. Plasmids

Plasmids pPKL4 (comprising the entire, functional fim gene
cluster) and pPKL114 (comprising the fimH gene) have been
30   described previously.

pPKL115 which is a plasmid containing the entire type 1 fim
gene cluster with a stop linker insertion in the fimH gene

(i.e. this plasmid expresses all the proteins necessary for the production of fimbriae except the FimH protein) was constructed in two steps:

(i) plasmid pPKL4 (refs. 27, 28) was digested with *KpnI* which
5 recognizes a unique restriction site in the *fimH* gene. The staggered end of the linearized plasmid was made blunt and ligated with the synthetic piece of DNA below (SEQ ID NO:7) containing stop codons in all three reading frames, resulting in plasmid pPKL114:

10   5'-GTCGACTTAATTAATTAAGTCGAC-3'
          3'-CAGCTGAATTAATTAATTCAGCTG-5';

(ii) the *HindIII-EagI* fragment from pPKL114, containing the entire *fim* gene cluster with the inactivated *fimH* gene was subsequently inserted into the *HindIII* and *EagI* sites of
15 plasmid pACYC184, resulting in plasmid pPKL115.

Plasmid pSM782 (generously provided by S. Molin, Department of Microbiology, Technical University of Denmark, DK-2800 Lyngby) containing the pre-S2 and S encoding regions of the hepatitis B viral genome, was made from plasmid λ-HBV1 (ref.
20 72) by subcloning a *EcoRI-DraI* fragment into pBR322.

Plasmid pLPA22 was constructed by inserting a 1018 bp *PvuII-MluI* fragment containing the *fimH* gene from pPKL4 into plasmid pUC18. The insert was positioned downstream and in a expression compatible orientation to the *lac* promoter resid-
25 ing on the vector part of the plasmid (Fig. 8). Expression in *E. coli* HB101 cells of functional FimH protein was monitored by complementing pLPA22 in trans with pPKL115 and testing for MS adhesion upon induction with IPTG.

Plasmids pLPA29 and pLPA30 were made by inserting 9-mer
30 asymmetric *BglII*-linkers into the *BsaAI* and *HincII* sites, respectively, in the *fimH* gene of plasmid pLPA22. At six different positions in the pLPA22 *fimH* gene a *BglII* site was

introduced without changing the reading frame, resulting in
plasmids pLPA98, pLPA36, pLPA58, pLPA30, pLPA29 and pLPA59
(Fig. 10). This was done either by inserting a *BglII* linker
into an appropriately treated restriction enzyme site, or by
5   changing 1-3 basepairs using PCR and thereby creating a *BglII*
site.

The plasmid pLPA36 was prepared by opening the pLPA22 *fimH*
gene with the restriction enzyme *Tth111I* and making the ends
blunt using Klenow polymerase and ligating using an 8 mer
10 *BglII* linker (SEQ ID NO:8):

        5'-CAGATCTG-3'
        3'-GTCTAGAC-5'

Plasmids pLPA58 and pLPA59 were made by *BglII* site-creating
site-directed mutagenesis of pLPA22 using standard PCR and
15 plasmid pLPA98 was constructed by opening the *fimH* gene,
making the ends blunt with T4 DNA polymerase and ligating
with the below 10 mer *BglII* linker (SEQ ID NO:9):

        5'-GAAGATCTTC-3'
        3'-CTTCTAGAAG-5'

20 Of the six resulting mutated *fimH* genes, three expressed
protein that was integrated into type 1 fimbriae, and at the
same time exhibited mannose-sensitive adhesion. Of these
three mutated FimH proteins, the two that conferred to *E.
coli* cells the strongest mannose-sensitive adhesion were
25 expressed from plasmids pLPA29 and pLPA30 (Fig. 9) and these
two plasmids were investigated further for their ability to
contain large mutations and still be biological active.

Plasmid pLPA29 has a 9 bp long symmetrical *BglII* linker
inserted into the *BsaAI* site 66 bp upstream of the stop codon
30 for the *fimH* gene, while plasmid pLPA30 has the same 9 bp
BglII linker inserted into the *HincII* site 163 bp upstream of
the stop codon of the *fimH* gene.

The plasmids pLPA37 and pLPA38 (Fig. 8) were constructed by inserting a 162 bp DNA fragment encoding the pre-S2 region of the Hepatitis B virus surface antigen into the BglII sites in pLPA29 and pLPA30, respectively. This DNA fragment was crea-

5 ted by a standard polymerase chain reaction (PCR) using the synthetic primers: (i) 5'-GGAGATCTAATTCCACAACCTT-3' (SEQ ID NO:11) and (ii) 5'-GGAGATCTGTTCAGCGCAGGGT-3' (SEQ ID NO:12), and plasmid pSM782 as a template.

A fragment of plasmid pLPA38 comprising the inserted
10 heterologous sequence encoding the pre-S2 region of hepatitis B surface antigen is shown in the below table wherein the heterologous sequence is underlined and the numbers indicated correspond to the positions of the amino acid residues in the mature FimH protein.

```
                    Bgl II
CAG TTC AGA TCT AAT TCC ACA ACC TTC CAC CAA ACT CTG CAA GAT
Gln Phe Arg Ser Asn Ser Thr Thr Phe His Gln Thr Leu Gln Asp
224          _____


CCC AGA GTG AGA GGC CTG TAT TTC CCT GCT GGT GGC TCC AGT TCA
Pro Arg Val Arg Gly Leu Tyr Phe Pro Ala Gly Gly Ser Ser Ser
_____


GGA ACA GTA AAC CCT GTT CTG ACT ACT GCC TCT CCC TTA TCG TCA
Gly Thr Val Asn Pro Val Leu Thr Thr Ala Ser Pro Leu Ser Ser
_____


                                              Bgl II
ATC TTC TCG AGG ATT GGG GAC CCT GCG CTG AAC AGA TCT TCG ACG
Ile Phe Ser Arg Ile Gly Asp Pro Ala Leu Asn Arg Ser Ser Thr
_____       226
```

15 The plasmids pLPA95 and pLPA93 (Fig. 8) were then made by inserting the below 51 bp synthetic double stranded DNA segment encoding amino acids 50-64 (comprising an epitope) of the B subunit of the cholera toxin into the BglI sites on pLPA30 and pLPA29, respectively (SEQ ID NO:10):

5'-GATCTGTTGAAGTTCCGGGTAGTCAGCATATCGATAGTCAGAAAAAAGCTG   -3'
3'-   ACAACTTCAAGGCCCATCAGTCGTATAGCTATCAGTCTTTTTTCGACCTAG-5'

A fragment of plasmid pLPA93 comprising the heterologous
sequence encoding the above DNA segment of the B subunit of
5   the cholera toxin is shown in the below table wherein the
heterologous sequence is underlined and the numbers indicated
correspond to the positions of the amino acid residues in the
mature FimH protein.

                 Bgl II
CAG TTC AGA TCT GTT GAA GTT CCG GGT AGT CAG CAT ATC GAT AGT
Gln Phe Arg Ser Val Glu Val Pro Gly Ser Gln His Ile Asp Ser
224   _____

                        BamH I/*Bgl* II
CAG AAA AAA GCT GGA TCT TCG ACG
Gln Lys Lys Ala Gly Ser Ser Thr
_____ 226

3.1.3. DNA techniques

10   Isolation of plasmid DNA was carried out according to the
method of Birnboim and Doly (ref. 73). Restriction endonu-
cleases were used according to the manufacturer's specifica-
tions (Biolabs). DNA sequencing was carried out by the di-
deoxy chain termination technique (ref. 49) using a sequenase
15   version 2.0 kit from USB. Oligonucleotides were made at the
core facilities of the Department of Microbiology, Technical
University of Denmark.

3.1.4. PCR methodology

Polymerase chain reactions (PCR) were performed on a Perkin
20   Elmer Cetus DNA Thermal Cycler 480. Reactions were set up as
100 μl volumes containing 200 μM each of dATP, dCTP, dGTP and
dTTP, 0.2-1.0 μM of each of the two primers, 2 mM MgCl$_2$, 10
mM Tris-HCl (pH 8.3), 50 mM KCl, 2.5 units of AmpliTaq DNA

polymerase and 0.1-0.2 µg of plasmid template. The reactions
were run for 25-30 cycles each consisting of 1 min. at 94°C,
1 min. at 40°C, and 1 min. at 72°C. For amplification of the
pre-S2 fragment the above primers 5'GGAGATCTAATTCCACAACCTT 3'
5   (SEQ ID NO:11) and 5'GGAGATCTGTTCAGCGCAGGGT 3'(SEQ ID NO:12)
were used.

3.1.5. Hemagglutination

The capacity of bacteria to express a D-mannose binding
phenotype was assayed by their ability to agglutinate guinea
10 pig erythrocytes on glass slides. Aliquots of liquid cultures
grown to an optical density of 3.0 and 5% erythrocytes were
mixed, and the time until agglutination occurred was
measured.

3.1.6. Antisera

15 Rabbit anti-type 1 fimbria serum raised against purified type
1 fimbriae has previously been described (ref. 74). A
monoclonal antibody directed against FimH (ref. 75) was
kindly provided from Dr. Maryvonne Dho-Moulin, Institut
National de la Recherche Agronomique, France. Goat serum
20 raised against cholera toxin B subunit (international stan-
dard for WHO No. 12-246) produced at the State Serum Insti-
tute, Copenhagen, Denmark was kindly provided by same insti-
tute. A monoclonal antibody directed against the pre-S2
domain of Hepatitis B surface antigen (ref. 76) was kindly
25 provided by Dr. Makoto Mayumi, Jichi Medical School, Japan.
Fluorescein (FITC) conjugated anti rabbit, anti mouse, or
anti goat sera were provided from Dakopats, Denmark.

3.1.7. Fluorescence labelling and CCD microscopy

Cells from overnight cultures (IPTG-induced, if required)
30 were harvested, washed in PBS and fixed for 10 minutes at
room temperature in a 3.5% (w/v) solution of paraformaldehyde
in PBS. Samples of 20 µl were placed on a poly-L-lysine

coated slide and air dried. After washing in PBS, 16 μl of a
1:5 (monoclonal) or 1:25 (polyclonal) dilution of the primary
antiserum was placed on top of each sample and left in a
moist incubation chamber for 1 hour. The slides were washed
5 three times in PBS and 16 μl of FITC conjugated antiserum
were added. After two hours in the dark, the slides were
washed three times in PBS and a drop of Citiflour (Citiflour
Ltd., London, U.K.) was placed on top of each sample. For
visualization, a Carl Zeiss Axioplan microscope equipped for
10 epifluorescence and phase-contrast was employed. Using a
charge-coupled device (CCD) camera, pictures were captured as
12-bit files with PMIS software (Photometrics) and
subsequently transferred to a Macintosh Quadra 950 computer
for image analysis.

15 3.1.8. Electron microscopy.

Electron microscopy and immuno-electron microscopy was
carried out essentially as described previously (ref. 61). In
brief, a 25 μl aliquot of bacterial suspension was placed on
a carbon-coated, glow discharged grid for 30 seconds. Grids
20 were washed in 2 drops of PBS, dehydrated for 5 min in each
of the following concentrations of ethanol: 25%, 50%, 75% and
96%, blotted dry and shadowed with tungsten wire at an angle
of 30°. For immuno-electron microscopy a monoclonal antibody
directed against the pre-S2 region was used diluted 1:5 as
25 the primary antibody and rabbit anti-mouse serum conjugated
with 10 nm gold particles (Dako) was used in dilution 1:20 as
the secondary antibody.

3.2. Results

As described above, two positions in the C-terminal part of
30 the FimH protein were engineered to contain heterologous
sequences mimicing foreign antigenic determinants. In the
present study, double plasmid systems were used. In each

plasmid pair one encoded either a wild-type or an engineered
version of the *fimH* gene, whereas the second plasmid encoded
auxiliary functions such as the two-component Fim-specific
transport system, regulatory genes and other structural
5   components of the fimbrial organelle except FimH (Table 3).

3.2.1. Engineering new restriction sites into *fimH*.

Based on algorithms for prediction of such parameters as
hydrophilicity and secondary structure, two potentially
10 optimal positions for insertions of heterologous sequences in
the C-terminal domain of the FimH protein were selected.
These correspond to positions 225 and 258 in the mature
protein predicted to be situated in a surface-exposed part of
the FimH protein. In order to facilitate later manipulations,
15 the *fimH* gene was subcloned into the pUC18 vector resulting
in plasmid pLPA22. Subsequently a *Bgl*II site was introduced
in-frame into positions 225 and 258, respectively. This was
carried out by site-directed mutagenesis employing synthetic
oligomers resulting in plasmids pLPA30 and pLPA29, respect-
20 ively (Fig. 9).

The introduced *Bgl*II sites resulted in a codon change from a
Leu to a Phe codon in position 225 and addition of codons for
the sequence Arg-Ser-Ser, in the case of plasmid pLPA29, and
addition of codons for the sequence Arg-Ser-Gly after posi-
25 tion 258 in the case of plasmid pLPA30. Sequence analysis of
the entire modified *fimH* genes in plasmids pLPA29 and pLPA30
confirmed that no other changes had occurred. Host cells
which in addition to plasmid pLPA29 or pLPA30 also contained
plasmid pPKL115 (*fimH*), showed wild-type phenotypic charac-
30 teristics with regard to adhesion and fimbriation as judged
by such criteria as hemagglutination (Table 3) and immuno-
fluorescence microscopy.

3.2.3. Engineering heterologous DNA-sequences encoding the
pre-S2 domain of hepatitis B surface antigen and a cholera
toxin epitope into _fimH_.

       As heterologous reporter epitopes the pre-S2 region of the
 5   hepatitis B surface antigen and a well characterized region
       of the B subunit of cholera toxin were selected. The pre-S2
       region is known to contain immunologically important (and
       protective) antigenic determinants (ref. 76). In addition,
       this region is disulphide bond-independent and apparently
10   more immunogenic than the major S protein. The cholera toxin
       segment consists of residues 50-64 of the B subunit and has
       previously been shown to elicit antibodies that bind to and
       neutralize cholera toxin (ref. 77).

15   A DNA segment of 162 nucleotides encoding 52 of the 55 amino
       acids of the pre-S2 region was amplified by PCR technology
       using plasmid pSM782 as template and primers that provided
       the amplified sequence with flanking _Bgl_II sites. Following
       restriction with _Bgl_II and purification the amplified frag-
20   ment was inserted into the BglII sites of plasmids pLPA29 and
       pLPA30 resulting in plasmids pLPA37 and pLPA38, respectively
       (Fig 9). Subsequent sequence analysis confirmed that the
       inserts were correctly oriented and that the reading frame of
       the chimeric _fimH_-pre-S2 genes was correct.
25

       A synthetic DNA segment encoding the cholera epitope was made
       by annealing two complementary 51 bp oligonucleotides which
       were designed to result in a double stranded DNA fragment
       with a _Bgl_II overhang in one end, a _Bam_HI overhang in the
30   other and an internal _Cla_I site. The epitope-encoding segment
       was inserted into the _Bgl_II site in the _fimH_ gene in plasmids
       pLPA29 and pLPA30, resulting in regeneration of a _Bgl_II site
       at only one end of the insert. This feature was used to
       identify plasmids with correct orientation of the insert. The
35   presence of the _Cla_I site was used for initial screening for
       clones containing the insert. Sequence analysis of plasmid
       pLPA93 and pLPA95, both harbouring the epitope-encoding

segment confirmed the orientation and conservation of the
reading frame in the chimeric *fimH*-cholera genes (Fig 8).

3.2.4. Expression of chimeric FimH adhesin comprising as

5 heterologous sequences the pre-S2 domain of hepatitis B
surface antigen and a cholera toxin epitope.

To evaluate whether the heterologous inserts in *fimH* were
compatible with protein expression the T7 polymerase/promoter
system of Tabor and Richardson (ref. 78) was used. Subcloning
10 into the pGEM3 vector system and subsequent assaying revealed
that proteins with the expected sizes were produced in all
cases from the chimeric *fimH* genes. More importantly, to
assess whether the FimH proteins harbouring foreign inserts
were accepted by the type 1 fimbrial transport system and
15 additionally, whether they were present on the bacterial
surface in a biologically functional form, the adhesion
phenotype of recombinant strains expressing the chimeric FimH
proteins was studied.

Bacterial hosts which in addition to plasmid pLPA38 (pre-S2
20 insert in position 225 in FimH) also contained plasmid
pPKL115 (*fimH*) gave, when induced by IPTG, good agglutination
of guinea-pig erythrocytes indicating the presence of a
biologically active form of the FimH adhesin on the cells
(Table 3). The combination of plasmids pLPA37 (pre-S2 in
25 position 258 in FimH) and pPKL115 resulted in weaker, but
detectable, hemagglutination (Table 3). Furthermore, such
cells were also shown by electron microscopy to have essen-
tially normal fimbriation (Fig. 10).

Table 3. Genotype and phenotype of plasmids (A, B or U, respectively indicate pACYC184, pBR322 or pUC18 based vector) used in this study, position of inserts and hemagglutination titer. Host cell: _E. coli_ HB101

5

Plasmid	relevant genotype	insert position	hemagglu- tination[a]
pPKL4 (B)	all _fim_ genes		15
pPKL115 (A)	_fimH_		>600
10 pLPA22 (U)	_fimH_+		>600
pLPA29 (U)	_fimH_-BglII	258	>600
pLPA30 (U)	_fimH_-BglII	225	>600
pLPA37 (U)	_fimH_-pre-S2	258	>600
pLPA38 (U)	_fimH_-pre-S2	225	>600
15 pLPA93 (U)	_fimH_-cholera	225	>600
pLPA95 (U)	_fimH_-cholera	258	>600
pLPA22 (U) +pPKL115 (A)	_fimH_+ _fimH_		10
20 pLPA29 +pPKL115	_fimH_-BglII _fimH_		7
pLPA30 +pPKL115	_fimH_-BglII _fimH_		8
pLPA37 +pPKL115	_fimH_-pre-S2 _fimH_		210
25 pLPA38 +pPKL115	_fimH_-pre-S2 _fimH_		100
pLPA93 +pPKL115	_fimH_-cholera _fimH_		11
30 pLPA95 +pPKL115	_fimH_-cholera _fimH_		16

a) Hemagglutination of guinea-pig erythrocytes indicated in seconds before reaction occurred. The average values of 4 measurements are given.

35   In the cases where a sequence mimicing a cholera epitope had been inserted into FimH, viz. pLPA93 (insert in position 225)

and pLPA95 (insert in position 258), respectively, an agglu-
tination phenotype also resulted when either of these plas-
mids were complemented by plasmid pPKL115 (*fimH*) (Table 3).
Again, this suggested that in spite of the presence of
5   foreign peptide segments the chimeric FimH proteins were
still able to reach the bacterial surface and maintain its
adhesive function. In addition to the adherence phenotypes of
the various clones the presence of engineered FimH adhesins
on the surface of the cells were monitored by CCD microscopy
10   in connection with fluorescent antibody methodology employing
a FimH-specific monoclonal serum. In all cases, significant
signals, albeit of varying intensity, were detected when
compared to a negative control strain that harboured the
auxiliary plasmid, pPKL115, alone.

15   3.2.5. Immunological detection of the pre-S2 segment of
hepatitis B surface antigen and the cholera toxin epitope in
chimeric FimH adhesins.

Since there was good evidence that the chimeric FimH proteins
were present on the surface of the *E. coli* hosts the ability
20   of specific antisera, raised against the pre-S2 part of
hepatitis B surface antigen or the cholera toxin B chain,
respectively to recognize the chimeric FimH-pre-S2 and FimH-
cholera proteins directly on the surface of the recombinant
bacteria were tested. By immunofluorescence microscopy *E.*
25   *coli* hosts harbouring either of plasmids pLPA37 or pLPA38 in
addition to plasmid pPKL115 were shown to react specifically
with antisera directed against the inserted heterologous
sequence, whereas hosts expressing wild-type FimH did not.
Similar results were obtained with the cholera toxin insert
30   in the same positions (plasmids pLPA93/pPKL115 and
pLPA95/pPKL115). Again, the heterologous inserts in the
chimeric FimH proteins were recognized by insert-specific
serum on the bacterial surface, whereas the relevant control
did not react.

35

These findings demonstrate that the foreign epitopes are
exposed on the surface of extracellularily located chimeric
FimH proteins and, significantly, in a conformation which
mimics the natural conformation of the epitope(s) as it
5   appears in the native hepatitis B surface antigen or the
native cholera toxin.

The results obtained by immunofluorescence microscopy were
corroborated by immuno-electron microscopy, employing the
pre-S2 specific monoclonal antibody as primary serum and a
10 colloid gold-labelled secondary antiserum. A significant
amount of gold particles were seen, mostly in connection with
the fimbrial organelles, on bacterial hosts harbouring chim-
eric *fimH*-pre-S2 genes (Fig. 10b and 10c), whereas only few
goldparticles were present on the control strain expressing
15 wild-type fimbriae (Fig. 10a). Furthermore, in the latter
case the gold-particles were not seen to be associated with
the fimbriae.

The plasmids pLPA22, pLPA29, pLPA30, pLPA37, pLPA38, pLPA93,
pLPA95 and pPKL115 in *E. coli* HB101 were deposited on 26
20 January 1994 with DSM, the Deutsche Sammlung von Mikro-
organismen und Zellkulturen GmbH, (German Collection of
Microorganisms and Cell Cultures), Mascheroder Weg 1B, D-
38124 Braunschweig, Germany, under the accession numbers DSM
8915, DSM 8916, DSM 8917, DSM 8918, DSM 8919, DSM 8920, DSM
25 8921 and DSM 8923, respectively.

EXAMPLE 4

Binding of the MFP class adhesin of *E. coli* CSH-50 to syn-
thetic peptides

14 synthetic peptides were synthesized on an ABI automated
30 peptide synthesizer according to the method of Merrifield
(Merrifield, R.B. 1963. Solid phase peptide synthesis.I. The
synthesis of tetrapeptide. J. Am. Chem. Soc. 85:2149). The

binding of the *E. coli* strain CSH-50 to these peptides were
tested essentially as described in Example 1. The results of
these binding assays indicated that this MFP class strain
adhered strongly to one group of peptides whereas the binding
5   the an other group of peptides was absent or weak. In the
below listing the one-letter code sequences of the synthetic
peptides are shown in a + group, i.e. the group of peptides
to which the tested strain adhered strongly, and a - group of
peptides to which the binding was weak or absent:

10   <u>+ group of peptides</u>

       FnSP1:   EAQQMVQPQSPVAVSQSKPGCYDNGKHYQI (SEQ ID NO:13)
       CB-II-G: EEGKRGARGEBGAAGPVGPBGERGARGNR (SEQ ID NO:14)
       sM1(19-32):   AIQNIRLRHENKDL (SEQ ID NO:15)
       sM6(1-11):   RVFPRGTVENPC (SEQ ID NO:16)
15   sM12(1-12):   DHSDLVAEKQRLC (SEQ ID NO:17)
       sM12(7-18):   AEKQRLEDLGQKC (SEQ ID NO:18)
       sM5(175-184):   TVKDKLAKEQC (SEQ ID NO:19)
       sM5(28-54):   KTKNEGLKTENEGLKTENEGLKTENEGC (SEQ ID NO:20)

       <u>- group of peptides</u>

20   sM5(134-163): QESKENEKALNELLEKTVKDKIAKEQENKE (SEQ ID NO:21)
       sM5(117-146): DLTKELNKTRQELANKQQESKENEKALNEL (SEQ ID NO:22)
       sM5(14-26):   KEALDKYELENHD (SEQ ID NO:23)
       sM6(22-31):   DVENSMLQAN (SEQ ID NO:24)
       sM5(55-84):   LKTEKSNLERKTAELTSEKKEHEAENDKLKC (SEQ ID NO:25)
25   sM24(289-303):HQKLEEQNKTSEASRC (SEQ ID NO:26)

       EXAMPLE 5

       <u>FimH adhesin of further clinical isolates</u>

       The following clinical isolates of *E. coli* were tested for
       adhesion class according to the methods described in Example
30   1: KB-23, KS-54, U221-3, MJ#9-3, MJ#31-3, MJ#11-2, MJ#2-2.

The results of these experiments are illustrated in Fig. 5.
As explained above, the isolate KB-23 showed the M^L type of
adhesion, and the isolate U221-3 expressed a M class adhesin
showing a mannose-resistant type of adhesion and accordingly,

5   this strain was classified as having a M^R class adhesin.
The amino acid sequences of these clinical isolates are shown
in Fig. 5 and their nucleotide sequences in Table 5 below.

Table 5 shows the nucleotide sequences of the *fimH* genes of
selected *fimH* genes disclosed in Example 1 [CI#3 (SEQ ID

10 NO:50), CI#4 (SEQ ID NO:44), CI#7 (SEQ ID NO:51), CI#10 (SEQ
ID NO:48) and CI#12 (SEQ ID NO:54)] and as the reference that
of the *E. coli* K12 strain PC31 as it was originally disclosed
by Klemm et al. (ref. 27) as the top sequence designated
PC31a and the sequence as it was determined recently (PC31b).

15 Additionally, the nucleotide sequences of the following
clinical isolates of *E. coli* are shown: KS54 (SEQ ID NO:52),
U221-3 (SEQ ID NO:53), MJ#9-3 (SEQ ID NO:46), MJ#31-3 (SEQ ID
NO:47), MJ#11-2 (SEQ ID NO:43), MJ#2-2 (SEQ ID NO:45) and F-
18 (SEQ ID NO:42).

Table 5. Nucleotide sequences of the above *fimH* genes disclosed in Example 1, *E. coli* K12 strain PC31 (PC31a and PC31b) and the nucleotide sequences of KS54, U221-3, MJ#9-3, MJ#31-3, MJ#11-2, MJ#2-2 and F-18.

```
        1094
PC31a   ATG AAA CGA GTT ATT ACC CTG TTT GCT GTA CTG CTG ATG GGC TGG TCG GTA AAT
F-18    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 4    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ22    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ9-3   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ31-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --- --- --- -A- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- --- --- -A- --- --- --- --- --- --- --- --- --- --- --- ---
KS54    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

        1148
PC31a   GCC TGG TCA TTC GCC TGT AAA ACC GCC AAT GGT ACC GCT ATC CCT ATT GGC GGT
F-18    --- --- --- --- --- --- --- --- --- --- --- --- --A --- --- --- --- ---
MJ11-2  --- --- --- --- --- --- --- --- --- --- --- --- --A --- --- --- --- ---
CI 4    --- --- --- --- --- --- --- --- --- --- --- --- --A --- --- --- --- ---
MJ22    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ9-3   --- --- --- --- --- --- --- --- --- --- --- --- --A --- --- --- --- ---
MJ31-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --T --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --T --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- --C --- --- --- --- --- ---
KS54    --- --- --- --- --- --- --- --- --- --- --- --C --- --- --- --- --- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

        1202
PC31a   GGC AGC GCC AAT GTT TAT GTA AAC CTT GCG CCC GTC GTG AAT GTG GGG CAA AAC
F-18    --- --- --- --- --- --- --- --- --- --- --T -C- --- --- --- --- --- ---
MJ11-2  --- --- --- --- --- --- --- --- --- --- --T -C- --- --- --- --- --- ---
CI 4    --- --- --- --- --- --- --- --- --- --- --T -C- --- --- --- --- --- ---
MJ22    --- --- --- --- --- --- --- --- --- --- --T -C- --- --- --- --- --- ---
MJ9-3   --- --- --T --- --- --- --- --- --- --- --T -C- --- --- --- --- --- ---
MJ31-3  --- --- --T --- --- --- --- --- --- --- --T -C- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- -C- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- -C- --- --- --- --- --- ---
CI 3    --- --- --- --- --- --- --- --- --- --- --- -C- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- -C- --- --- --- --- --- ---
KS54    --- --- --T --- --- --- --- --- --- --- --T -C- --- --- --- --- --- C--
U221-3  --- --- --- --- --- --- --- --- --- --- --- -C- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --- --T -C- --- --- --- --- --- ---
```

Table 5, continued

```
      1256
PC31a  CTG GTC GTG GAT CTT TCG ACG CAA ATC TTT TGC CAT AAC GAT TAT CCG GAA ACC
F-18   --- --- --A --- --- --- --- --- --- --- --- --- --- --- --C --A --- ---
MJ11-2 --- --- --A --- --- --- --- --- --- --- --- --- --- --- --C --A --- ---
CI 4   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --A --- ---
MJ22   --- --- --A --- --- --- --- --- --- --- --- --- --- --- --C --- --- ---
MJ9-3  --- --- --A --- --- --- --- --- --- --- --- --- --- --- --C --A --- ---
MJ31-3 --- --- --A --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 10  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7   --- --- --A --- --- --- --- --- --- --- --- --- --- --- --C --- --- ---
KS54   --- --- --A --- --- --- --- --- --- --- --- --- --- --- --C --- --- ---
U221-3 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- --- ---

      1312
PC31a  ATT ACA GAC TAT GTC ACA CTG CAA CGA GGC TCG GCT TAT GGC GGC GTG TTA TCT
F-18   --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- --- ---
MJ11-2 --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- --- ---
CI 4   --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- --- ---
MJ22   --- --- --- --- --- --- --- --- --- --T --- --- --- --- -A- --- --- ---
MJ9-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- -A- --- --- ---
MJ31-3 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 10  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3   --- --- --T --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7   --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- --- ---
KS54   --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- --- ---
U221-3 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12  --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- --- ---

      1366
PC31a  AAT TTT TCC GGG ACC GTA AAA TAT AGT GGC AGT AGC TAT CCA TTT CCT ACC ACC
F-18   -G- --- --- --- --- --- --- --- -A- --- --- --- --- --T --C --- --T ---
MJ11-2 -G- --- --- --- --- --- --- --- -A- --- --- --- --- --T --C --- --T ---
CI 4   -G- --- --- -A- --- --- --- --- -A- --- --- --- --- --T --C --- --T ---
MJ22   -G- --- --- --- --- --- --- --- -A- --- --- --- --- --T --C --- --T ---
MJ9-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --T --C --- --T ---
MJ31-3 --- --- --- --- --- --- --- --- --- --- --- --- --- --C --G --T ---
CI 10  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --G --T ---
PC31b  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --G --- ---
KS54   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
U221-3 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --G --T ---
CI 12  -G- --- --- --- --- --- --- --- -A- --- --- --- --- --T --C --- --T ---
```

Table 5, continued

```
          1420
PC31a   AGC GAA ACG CCG CGC GTT GTT TAT AAT TCG AGA ACG GAT AAG CCG TGG CCG GTG
F-18    --- --- --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --- --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 4    --- --- --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ22    --- --- --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ9-3   --- --- --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ31-3  --- --- --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --T --- --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- -T- --G --- --- --- --- --- --- --- --- --- --- --- --- ---
KS54    --- --- --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---


          1474
PC31a   GCG CTT TAT TTG ACG CCT GTG AGC AGT GCG GGC GGG GTG GCG ATT AAA GCT GGC
F-18    --- --- --- --- --- --G --- --- --- --- --G --A --- --- --- --- --- ---
MJ11-2  --- --- --- --- --- --G --- --- --- --- --G --A --- --- --- --- --- ---
CI 4    --- --- --- --- --- --- --- --- --- --- --G --A --- --- --- --- --- ---
MJ22    --- --- --- --- --- --G --- --- --- --- --G --A --- --- --- --- --- ---
MJ9-3   --- --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- ---
MJ31-3  --- --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --                -- --- --- ---
CI 3    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- --- --- -T- --- --- --- ---
KS54    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
U221-3  --- -G- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --G --A --- --- --- --- --- --- ---


          1528
PC31a   TCA TTA ATT GCC GTG CTT ATT TTG CGA CAG ACC AAC AAC TAT AAC AGC GAT GAT
F-18    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 4    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ22    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ9-3   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ31-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
KS54    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
```

Table 5, continued

```
        1582
PC31a   TTC CAG TTT GTG TGG AAT ATT TAC GCC AAT AAT GAT GTG GTG GTG CCT ACT GGC
F-18    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
CI 4    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
MJ22    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
MJ9-3   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
MJ31-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
KS54    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---


        1636
PC31a   GGC TGC GAT GTT TCT GCT CGT GAT GTC ACC GTT ACT CTG CCG GAC TAC CGT GGT
F-18    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -C- ---
MJ11-2  --- --T --- -C- --- --- --- --- --- --- --- --- --- --- --- --- -C- ---
CI 4    --- --T --- --- --- --- --- --- --- --- --- --- --- --- --- --- -C- ---
MJ22    --- --T --- --C --- --- --- --- --- --- --- T-- --- --- --- --- -C- ---
MJ9-3   --- --- --- --- --- --- -A- --- --- --- --- --- --- --- --- --- -C- ---
MJ31-3  --- --- --- --- --- --- -A- --- --- --- --- --- --- --- --- --- -C- ---
CI 10   --- --- --- --- --- --- --G --- --- --- --- --- --- --- --- --- -C- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -C- ---
CI 3    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -C- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -C- ---
KS54    --- --- --- --- --- --- -A- --- --- --- --- --- --- --- --- --- -C- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -C- ---
CI 12   --- --T --- --- --- --- T-- --- --- --- --- --- T-- --- --- --- -CA ---


        1690
PC31a   TCA GTG CCA ATT CCT CTT ACC GTT TAT TGT GCG AAA AGC CAA AAC CTG GGG TAT
F-18    --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 4    --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ22    --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ9-3   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ31-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
KS54    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
```

Table 5, continued

```
        1744
PC31a   TAC CTC TCC GGC ACA CAC GCA GAT GCG GGC AAC TCG ATT TTC ACC AAT ACC GCG
F-18    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --A --- --- --- --T --- --- --- --- --- --- --- --- --- --- --- ---
CI 4    --- --- --- --- --- G-- --- --- --- --- --- --- --- --- --- --- --- ---
MJ22    --- --A --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ9-3   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ31-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
KS54    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

        1798
PC31a   TCG TTT TCA CCT GCA CAG GGC GTC GGC GTA CAG TTG ACG CGC AAC GGT ACG ATT
F-18    --- --- --- --C --G --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --- --- --C --G --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 4    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ22    --- --- --- --C --G --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ9-3   --- --- --- --A --G --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ31-3  --- --- --- --A --G --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -A-
CI 7    --- --- --- --A --G --- --- --- --- --- --- --- --- --- --- --- --- ---
KS54    --- --- --- --A --G --- --- --- --- --- --- --- --- --- --- --- --- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

        1852
PC31a   ATT CCA GCG AAT AAC ACG GTA TCG TTA GGA GCA GTA GGG ACT TCG GCG GTG AGT
F-18    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A ---
CI 4    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
MJ22    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A ---
MJ9-3   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A ---
MJ31-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A ---
KS54    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A ---
CI 12   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --A ---
```

Table 5, continued

```
        1906
PC31a   CTG GGA TTA ACG GCA AAT TAT GCA CGT ACC GGA GGG CAG GTG ACT GCA GGG AAT
F-18    --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 4    --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- ---
MJ22    --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --C --- ---
MJ9-3   --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- ---
MJ31-3  --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- ---
CI 10   T-- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- ---
KS54    --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- ---

        1960
PC31a   GTG CAA TCG ATT ATT GGC GTG ACT TTT GTT TAT CAA
F-18    --- --- --- --- --- --- --- --- --- --- --- ---
MJ11-2  --- --- --- --- --- --- --- --- --- --- --- ---
CI 4    --- --- --- --- --- --- --- --- --- --- --- ---
MJ22    --- --- --- --- --- --- --- --- --- --- --- ---
MJ9-3   --- --- --- --- --- --- --- --- --- --- --- ---
MJ31-3  --- --- --- --- --- --- --- --- --- --- --- ---
CI 10   --- --- --- --- --- --- --- --- --- --- --- ---
PC31b   --- --- --- --- --- --- --- --- --- --- --- ---
CI 3    --- --- --- --- --- -C- --- --- --- --- --- ---
CI 7    --- --- --- --- --- --- --- --- --- --- --- ---
KS54    --- --- --- --- --- --- --- --- --- --- --- ---
U221-3  --- --- --- --- --- --- --- --- --- --- --- ---
CI 12   --- --- --- --- --- --- --- --- --- --- --- ---
```

EXAMPLE 6

Enrichment selection of strains having mutated FimH adhesins
conferring altered adhesion ability

One mechanism whereby new binding activities of bacterial
5   adhesins may arise is by random, naturally occurring mutage-
nesis. In nature, a variety of factors would enrich for
strains that possessed adhesive capacities conferring a
selective advantage. In the present example an in vitro
procedure was used to select for potential mutants with
10   altered adhesive capacity. As a target substratum bovine κ-
casein was selected.

κ-casein is the glycosylated isoform of bovine casein con-
sists of a single polypeptide chain containing 169 amino acid
residues the sequence of which has been determined (ref. 68).
15   Bovine κ-casein does not contain N-glycosidic linkages, but
up to six O-linked oligosaccharides are present in the C-
terminal region of the molecule (refs. 68, 69). The secchari-
de moieties are heterologous and also vary as a function of
time after parturition. Of significance for the present study
20   is the fact that D-mannose is not present in the bovine κ-
casein. Only di- to hexasaccharides containing galactose, N-
acetyl-galactosamine, N-acetyl-glucosamine, fucose and sialic
acid have been described (ref. 68). Glycoproteins having such
saccharide compositions would not be expected to serve as a
25   receptor for the classic type of the FimH adhesin such as is
found in E. coli strain PC31.

Adhesion tests were performed to verify the inability of
recombinant strains carrying the fimH gene from E. coli
strain PC31 to adhere to immobilized κ-casein. The E. coli
30   strain used, KB1001 is HB101 containing plasmids pPKL115 and
pLPA22 (ref. 70). The adhesion assay was performed using
microtiter plates coated with 30 μg/ml κ-casein in 0.1 M
sodium bicarbonate (pH 9.6) for 30 minutes, followed by
blocking any remaining binding sites with a subsequent in-

cubation with 0.1% bovine serum albumin (BSA) in PBS. A
quantitative adhesin assay was performed as described in more
detail elsewhere (ref. 71). Briefly, bacteria were diluted to
equivalent concentrations (5 x 10^7 cells/$100\mu l$) in PBS

5   containing 0.1% BSA, added to coated microtiter wells for 30
minutes at 37°C. After washing the wells thoroughly to remove
unbound bacteria, BHI broth was added and the bacteria were
allowed to grow at 37°C on a rotating platform (150 rpm)
until the optical density could be measured (2-2.5 hours).

10 Comparisons can be made of optical densities obtained in the
test wells to those obtained in standard curves developed
from the plating of known numbers of bacteria under similar
conditions, allowing extrapolation to absolute numbers of
bound bacteria (ref. 70).

15 The KB1001 strain comprising the *fimH* gene from PC31 bound to
immobilized mannan in significant numbers, but there was
substantially no measurable adhesion to immobilized κ-casein.
To select for possible mutant cells having acquired the
ability to bind to κ-casein, cells of KB1001 were allowed to

20 interact with κ-casein immobilized on microtiter wells. After
thorough washing to remove non-adhering bacterial cells,
cells adhering to the wells were collected and grown over-
night in BHI broth. These "enriched" bacterial cultures were
again allowed to interact with immobilized κ-casein, the

25 plates were washed and adhering cells collected in nutrient
broth. This enrichment cycle was repeated up to ten times.
Bacterial cells obtained from the last of these cycles ("en-
riched" strains) adhered to κ-casein in significantly in-
creased numbers in comparison to the parent ("non-enriched")

30 strain (Table 6.1). Individual colonies of "enriched" KB1001
were isolated and four tested for ability to adhere to κ-
casein. Three enriched cultures (clones) bound to κ-casein
significantly better than did the non-enriched parent strain.

Table 6.1. Adhesion to casein of non-enriched and enriched *E. coli* strain KB1001.

Strain	bacteria binding to κ-casein[a]
Non-enriched KB1001 (pPKL115 + pLPA22)	0.043 ± 0.018
Enriched KB1001 (pPKL115 + pLPA22)	0.249 ± 0.004

a) numbers represent optical density of bacterial growth $\pm$ S.D. with background O.D. substracted. N = 3.

To determine whether the new adhesive activity was due to plasmid-related changes and not simply to host cell-related changes, plasmid preparations of pLPA22 were made from enriched and from non-enriched strains and used to transform *E. coli* HB101 containing the auxiliary plasmid pPKL115. Randomly selected transformants resistant to ampicillin and chloramphenicol were tested for adhesion to κ-casein, and several of the transformants harbouring plasmids from enriched cultures adhered in significantly increased numbers relative to plasmid-containing cells of the non-enriched parent strain (Table 6.2).

Table 6.2. Adhesion to casein of HB101 (pPKL115) transformed with plasmids from enriched or non-enriched strain KB1001.

Plasmid derived from:	bacteria binding to κ-casein[a]
Non-enriched KB1001	$5 \pm 0.1 \times 10^3$
Enriched KB1001	$50 \pm 1.5 \times 10^3$

a) numbers represent mean number of bacteria per well ± S.D.
N = 3.

The above results demonstrate that random or spontaneous
mutations in genes coding for a bacterial adhesin that confer
binding to a new substratum (i.e. a receptor moiety to which
the parent strain does not bind), can be selected for by
appropriate in vitro procedures.

REFERENCES

1. Abraham, S.N., D. Sun, J.B. Dale and E.H. Beachey. 1988. Nature 336:682-684.

2. Abraham, S.N., D.L. Hasty, W.A. Simpson and E.H. Beachey. 1983. J. Exp. Med. 158:1114-1128.

3. Abraham, S.N., J.D. Goguen, D. Sun, P. Klemm and E. H. Beachey. 1987. J. Bacteriol. 169:5530-5536.

4. Abraham, S.N., J.P. Babu, C.S. Giampapa, D.L. Hasty, W.A. Simpson and E.H. Beachey. 1985. Infect. Immun 48:625-628.

5. Akiyama, S.K. and K.M. Yamada. 1985. J. Biol. Chem. 260:4492-4500.

6. Barondes, S.H. 1988. Trends in Biochem. Sci. 13:480-482.

7. Beachey, E.H. 1981. J. Infect. Dis. 143:325-345.

8. Blomfield, I.C., M.S. McClain and B.I. Eisenstein. 1991. Infect. Immun. 5: 1439-1445.

9. Bloch, C.A., B.A.D. Stocker and P. Orndorff. 1992. Mol. Microbiol. 6:697-701.

10. Brinton, C.C., Jr. 1959. Nature 183:782-786.

11. Brinton, C.C., Jr. 1965. Trans. N.Y. Acad. Sci. 27:1003-1005.

12. Dodd, D.C. and B.I. Eisenstein. 1982. Infect. Immun. 38:764-773.

13. Duguid, J.P. and D.C. Old. 1980. Adhesive properties of Enterobacteriaceae, p. 185-217. *In* E.H. Beachey (ed.), Bac-

terial Adherence (Receptors and Recognition Series B, vol. 6). Chapman and Hall, London.

14. Duguid, J.P., J.W. Smith, G. Dempster and P.N. Edmunds. 1955. J. Pathol. Bacteriol. 70:335-354.

5   15. Engvall, E. and E. Ruoslahti. 1977. J. Cancer 20:1-5.

16. Firon, N., I. Ofek and N. Sharon. 1983. Carbohydrate Res. 120:235- 249.

17. Gerlach, G.-F., S. Clegg and B.L. Allen. 1989. J. Bacteriol. 171:1262-1270.

10   18. Gibbons, R.J. 1984. J. Dent. Res. 63:378-385.

19. Guerina, N.G., T.W. Kessler, V.J. Guerina, M.R. Neutra, H.W. Clegg, S. Langerman, F.A. Scannapieco and D.A Goldman. 1983. J. Infect. Dis. 148:395-405.

20. Hanson, M.S., J. Hempel and C.C. Brinton, Jr. 1988. J.
15   Bacteriol. 170:3350-3358.

21. Hanson, M.S. and C.C. Brinton, Jr. 1988. Nature (London) 322:265-268.

22. Harris, S.L., D.A. Elliott, M.C. Blake, L.M. Must, M. Messenger and P.E. Orndorff. 1990. J. Bacteriol. 172:6411-
20   6418.

23. Hornick, D.B., B.L. Allen, M.A. Horn and S. Clegg. 1991. J. Clin. Microbiol. 29:1795-1800.

24. Hull, R.A., R.E. Gill, P. Hsu, B.H. Minshew and S. Falkow. 1981. Infect. Immun. 33: 933-938.

25   25. Hultgren, S.J., T.N. Porter, A.J. Schaeffer and J.L. Duncan. 1985. Infect. Immun. 50:370-377.

26. Keith, B., L. Maurer, P. Spears and P. Orndorff. 1986.
Infect Immun. 53:693-696.

27. Klemm, P. and G. Christensen. 1987. Mol. Gen. Genet.
208:439-445.

5   28. Klemm, P., B.J. Jorgensen, I. van Die, H. de Ree and H.
Bergmans. 1985. Mol. Gen. Genet. 199:410-414.

29. Krogfelt, K.A., H. Bergmans and P. Klemm. 1990. Infect.
Immun. 58: 1995-1998.

30. Krusius, T., M. Fukuda, A. Dell and E. Ruoslahti. 1985.
10 J. Biol. Chem. 260:4110-4116.

31. Lockman, H.A. and R. Curtiss, III. 1992. Infect. Immun.
60:491-496.

32. Maurer, L. and P. Orndorff. 1987. J. Bacteriol. 169:640-
645.

15 33. Mayaan, M., I. Ofek, O. Medalia and M. Aronson. 1985.
Infect. Immun. 49:785-789.

34. McCormick, B.A., D.P. Franklin, D.C. Laux and P.C. Cohen.
1989. Infect. Immun. 57:3022-3029.

35. Miller, J.H. 1972. Experiments In Molecular Genetics.
20 Cold Spring Harbor Press, Cold Spring Harbor, New York.

36. Minion, F.C., S.N. Abraham, E.H. Beachey and J.D. Goguen.
1986. J. Bacteriol. 165:1033-1036.

37. Neeser, J.-R., B. Koellreutter and P. Wuersch. 1986.
Infect. Immun. 52:428-436.

25 38. Nishikawa, A., T. Sekine, R. Ikeda, T. Shinoda and Y.
Fukazawa. 1990. Microbiol. Immunol. 34: 825-840.

39. Ohman, D.E. 1988. Experiments in Gene Manipulation. Prentice Hall, Inc., Englewood Cliffs, New Jersey.

40. Ofek, I. and N. Sharon. 1990. Curr. Top. Microbiol. Immunol. 151:91-113.

5 41. Ofek, I., D. Mirelman and N. Sharon. 1977. Nature 265:623-625.

42. Ofek, I. and E.H. Beachey. 1978. Infect. Immun. 22:247-254.

43. Old, D.C. 1972. J. Gen. Microbiol. 71: 149-157.

10 44. Orndorff, P.E. and S. Falkow. 1984. J. Bacteriol. 159:736-744.

45. Petersen, T.E., K. Skorstengaard and K. Vibe Pedersen. 1990. Primary structure of fibronectin, p. 1-24. In D. Mosher (ed.), Fibronectin. Academic Press, New York.

15 46. Quiocho, F.A. 1986. Ann Rev. Biochem. 55:287-315.

47. Rosenstein, I.J., M. Stoll, T. Mizuochi, R. Childs, E. Hounsell and T. Feizi. 1988. The Lancet ii:1327-1330.

48. Sambrook, J., E. Fritsch and T. Maniatis. 1989. Molecular Cloning. CSH Laboratory Press.

20 49. Sanger, F., S. Nicklen and A.R. Coulson. 1977. Proc. Natl. Acad. Sci. U.S.A. 74:5463-5467.

50. Sharon, N. and I. Ofek. 1986. Bacterial surface lectins specific for mannose, p. 55-81. In D. Mirelman (ed.), Microbial Lectins and Agglutinins, J. Wiley and Sons, New York.

25 51. Sokurenko, E.V., H.S. Courtney, S.N. Abraham, P. Klemm and D.L. Hasty. 1992. Infect. Immun. 60:4709-4719.

52. Strömberg, N., B.-I. Marklund, B. Lund, D. Ilver, A. Hamers, W. Gaastra, K-A. Karlsson and S. Normark. 1990. EMBO J. 9:2001-2010.

53. Strömberg, N., P-G. Nyholm, I. Pascher and S. Normark.
5   1991. Proc. Natl. Acad. Sci. U.S.A. 88:9340-9344.

54. Takasaki, S., K. Yamashita, K. Suzuki, S. Iwanaga and A. Kobata. 1979. J. Biol. Chem. 254:8548-8553.

55. Tewari, R., J.I. MacGregor, T. Ikeda, J.R. Little, S.C. Hultgren and S.N. Abraham. 1993. J. Biol. Chem. 268:3009-
10   3015.

56. Westerlund, B., P. Kuusela, T. Vartio, I. van Die and T. K. Korhonen. 1989. FEBS Letters 243:199-204.

57. Westerlund, B., I. van Die, C. Kramer, P. Kuusela, H. Holthöfer, A-M. Tarkkanen, R. Virkola, N. Riegman, H. Berg-
15   mans, W. Hoekstra and T.K. Korhonen. 1991. Mol. Microbiol. 5:2965-2975.

58. Zardi, L., B. Carnemolla, E. Balza, L. Borsi, P. Castellani, M. Rocco and A. Siri. 1985. Eur. J. Biochem. 146: 571-579.

20   59. Hultgren, S.J., F. Lindberg, G. Magnusson, J. Kihlberg, J.M. Tennent and S. Normark. 1989. Proc. Natl. Acad. Sci. U.S.A., 86:4357-4361.

60. Kaufman, M.R. and R.K. Taylor 1994. Fimbriae of *Vibrio cholerae*. *In* Fimbriae: Adhesion, Genetics, Biogenesis and
25   Vaccines. Ed. P. Klemm. CRC Press, Boca Raton, pp. 165-179.

61. Klemm, P. and K.A. Krogfelt. 1994. Type 1 fimbriae of *Escherichia coli*. *In* Fimbriae: Adhesion, Genetics, Biogenesis and Vaccines. Ed. P. Klemm. CRC Press, Boca Raton, pp. 9-28.

62. Krogfelt, K. A. 1991. Rev. Infect. Dis. 13:721-735.

63. Kuehn, M. J., D. Haslam, S. Normark and S. J. Hultgren. 1994. Structure, function and biogenesis of *Escherichia coli* P pili. *In* Fimbriae: Adhesion, Genetics, Biogenesis and Vaccines. Ed. P. Klemm. CRC Press, Boca Raton, pp. 39-55.

64. Schmidt, M.A. 1994. Non-fimbrial adhesins of *Escherichia coli*. In Fimbriae: Adhesion, Genetics, Biogenesis and Vaccines. Ed. P. Klemm. CRC Press, Boca Raton, pp. 89-100.

65. Sharon, N. 1987. FEBS Lett. 217:145-150.

66. Sokurenko, E.V., H.S. Courtney, D.E. Ohman, P. Klemm and D.L. Hasty. 1994. J. Bacteriol. 176:748-755.

67. Pedersen, J.C. and T.D. Leser. 1992. Microbial Releases. 1:95-102.

68. Jollès, J., F. Schoentgen, C. Alais, A.-M. Fiat and P. Jollès. 1972. Helv. Chim. Acta. 55:2872-2883.

69. Fiat, A.-M., C. Alais and P. Jollès. 1968. Chimia 22:137-139.

70. Pallesen, L., L.K. Poulsen, L.B. Jensen, G. Chistiansen and P. Klemm. 1995. Chimeric FimH adhesin of type 1 fimbriae as presenter of heterologous epitopes. J.Bacteriol. submitted for publication.

71. Sokurenko, E.V., V. McMackin and D.L. Hasty. 1994. Growth assay for bacterial adhesion to immobilized ligands. In R. Doyle and I. Ofek (eds), Methods in Enzymology. Vol 252, Methods for Microbial Adhesion. Academic Press, N.Y. in press.

72. Charnay, P., C. Pourcel, A. Louise, A. Fritsch and P. Tiollais. 1979. Proc. Natl. Acad. Sci. USA 76:2222-2226.

73. Birmboim, H.C. and J. Doly. 1979. Nucl. Acid Res. 7:1513-1523.

74. Krogfelt, K.A. and P. Klemm. 1988. Microb. Pathog. 4:231-238.

5   75. Chanteloup, N.K., M. Dho-Moulin, E. Esnault, A. Brée and J.-P. Lafont. 1991. Microbial Path. 10:271-280.

76. Itoh, Y., E. Takai, H. Ohnuma, K. Kitajama, F. Tsuda, A. Machida, S. Mishiro, T. Nakamura, Y. Miyakawa and M. Mayumi. 1986. Proc. Natl. Acad. Sci. USA 83:9174-9178.

10   77. Jacob, O.C., M. Sela and R. Arnon. 1983. Proc. Natl. Acad. Sci. USA. 80:7611-7615.

78. Tabor, S. and C. Richardson. 1985. Proc. Natl. Acad. Sci. USA. 82:1074-1078.

8 6

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 53 , line 26 .	

B. IDENTIFICATION OF DEPOSIT   Further deposits are identified on an additional sheet [X]

Name of depositary institution

DSM-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH

Address of depositary institution (*including postal code and country*)

Mascheroder Weg 1B
D-38124 Braunschweig
Germany

Date of deposit	Accession Number
26 January 1994	DSM 8922

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)   This information is continued on an additional sheet [ ]

As regards the respective Patent Offices of the respective desig-
nated states, the applicants request that a sample of the deposi-
ted microorganisms only be made available to an expert nominated
by the requester until the date on which the patent is granted or
the date on which the application has been refused or withdrawn or
is deemed to be withdrawn.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
[X] This sheet was received with the international application	[ ] This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

INDICATIONS RELATING TO DEPOSITED MICROORGANISMS
(PCT Rule 12*bis*)

Additional sheet

In addition to the microorganism indicated on page 53 of the
5 description, the following microorganisms have been deposited
with

        DSM-Deutsche Sammlung von Mikroorganismen und
        Cellkulturen GmbH
        Mascheroder Weg 1b, D-38124 Braunschweig, Germany

10 on the dates and under the accession numbers as stated below:

Accession number	Date of deposit	Description Page No.	Description Line No.
DSM 8915	26 January 1994	66	24
DSM 8916	26 January 1994	66	24
DSM 8917	26 January 1994	66	24
DSM 8918	26 January 1994	66	24
DSM 8919	26 January 1994	66	24
DSM 8920	26 January 1994	66	24
DSM 8921	26 January 1994	66	25
DSM 8923	26 January 1994	66	25

For all of the above-identified deposited microorganisms, the
following additional indications apply:

25 As regards the respective Patent Offices of the respective
designated states, the applicants request that a sample of
the deposited microorganisms stated above only be made
available to an expert nominated by the requester until the
date on which the patent is granted or the date on which the
30 application has been refused or withdrawn or is deemed to be
withdrawn.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT:
        (A) NAME: GX BioSystems A/S
        (B) STREET: Mothsvej 70
        (C) CITY: Holte
        (D) COUNTRY: Denmark
        (E) POSTAL CODE (ZIP): 2840

    (ii) TITLE OF INVENTION: Receptor specific bacterial adhesins and
        their use

    (iii) NUMBER OF SEQUENCES: 55

    (iv) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5               10              15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20              25              30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Val
        35              40              45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50              55              60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65              70              75              80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
            85              90              95

Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100             105             110
```

```
Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115                 120             125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130                 135             140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145                 150             155                 160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165             170             175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr
            180             185             190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
        195             200             205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210             215             220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225             230             235                 240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245             250             255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260             265             270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
        275             280             285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290             295             300
```

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Glu Ala Gln Gln Met Val Gln Pro Gln Ser Pro Val Ala Val Ser Gln
1               5               10                  15

Ser Lys Pro Gly Cys Tyr Asp Asn Gly Lys His Tyr Gln Ile
            20              25              30
```

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA (genomic)

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGGGGGTGCA CACCTACAGC TGAACCCGG   29

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA (genomic)

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGGGTGCAC TCAGGGAACC ATTCAGGCA   29

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA (genomic)

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGGTGCGCAT TATTGATAAA CAAAAGTCAC   30

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

    GGGGCATGCT TATTGATAAA CAAAAGTCAC   30

    (2) INFORMATION FOR SEQ ID NO:7:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

    GTCGACTTAA TTAATTAAGT CGAC   24

    (2) INFORMATION FOR SEQ ID NO:8:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 8 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

    CAGATCTG   8

    (2) INFORMATION FOR SEQ ID NO:9:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 10 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

    GAAGATCTTC   10

    (2) INFORMATION FOR SEQ ID NO:10:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 51 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single

        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GATCCAGCTT TTTTCTGACT ATCGATATGC TGACTACCCG GAACTTCAAC A   51

(2) INFORMATION FOR SEQ ID NO:11:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 22 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAGATCTAA TTCCACAACC TT   22

(2) INFORMATION FOR SEQ ID NO:12:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 22 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGAGATCTGT TCAGCGCAGG GT   22

(2) INFORMATION FOR SEQ ID NO:13:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: unknown
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Glu Ala Gln Gln Met Val Gln Pro Gln Ser Pro Val Ala Val Ser Gln
1   5   10   15

```
      Ser Lys Pro Gly Cys Tyr Asp Asn Gly Lys His Tyr Gln Ile
            20                      25                      30
```

(2) INFORMATION FOR SEQ ID NO:14:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 29 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: unknown
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
      Glu Glu Gly Lys Arg Gly Ala Arg Gly Glu Asx Gly Ala Ala Gly Pro
      1                   5                   10                      15

      Val Gly Pro Asx Gly Glu Arg Gly Ala Arg Gly Asn Arg
                  20                      25
```

(2) INFORMATION FOR SEQ ID NO:15:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 14 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: unknown
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
      Ala Ile Gln Asn Ile Arg Leu Arg His Glu Asn Lys Asp Leu
      1                   5                   10
```

(2) INFORMATION FOR SEQ ID NO:16:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 12 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: unknown
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
      Arg Val Phe Pro Arg Gly Thr Val Glu Asn Pro Cys
      1                   5                   10
```

(2) INFORMATION FOR SEQ ID NO:17:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 13 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS: unknown
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: peptide

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Asp His Ser Asp Leu Val Ala Glu Lys Gln Arg Leu Cys
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:18:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 13 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS: unknown
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: peptide

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Ala Glu Lys Gln Arg Leu Glu Asp Leu Gly Gln Lys Cys
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:19:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 11 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS: unknown
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: peptide

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Thr Val Lys Asp Lys Leu Ala Lys Glu Gln Cys
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:20:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 28 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS: unknown
      (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Lys Thr Lys Asn Glu Gly Leu Lys Thr Glu Asn Glu Gly Leu Lys Thr
1   5   10   15

Glu Asn Glu Gly Leu Lys Thr Glu Asn Glu Gly Cys
            20   25

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Gln Glu Ser Lys Glu Asn Glu Lys Ala Leu Asn Glu Leu Leu Glu Lys
1   5   10   15

Thr Val Lys Asp Lys Ile Ala Lys Glu Gln Glu Asn Lys Glu
            20   25   30

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Asp Leu Thr Lys Glu Leu Asn Lys Thr Arg Gln Glu Leu Ala Asn Lys
1   5   10   15

Gln Gln Glu Ser Lys Glu Asn Glu Lys Ala Leu Asn Glu Leu
            20   25   30

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Lys Glu Ala Leu Asp Lys Tyr Glu Leu Glu Asn His Asp
1   5   10

(2) INFORMATION FOR SEQ ID NO:24:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Asp Val Glu Asn Ser Met Leu Gln Ala Asn
1   5   10

(2) INFORMATION FOR SEQ ID NO:25:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Leu Lys Thr Glu Lys Ser Asn Leu Glu Arg Lys Thr Ala Glu Leu Thr
1   5   10   15

Ser Glu Lys Lys Glu His Glu Ala Glu Asn Asp Lys Leu Lys Cys
            20   25   30

(2) INFORMATION FOR SEQ ID NO:26:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
His Gln Lys Leu Glu Glu Gln Asn Lys Thr Ser Glu Ala Ser Arg Cys
1               5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO:27:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5                   10                  15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20              25              30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
        35              40              45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50              55              60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65              70              75              80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
            85              90              95

Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100             105             110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115             120             125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130             135             140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145             150             155             160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165             170             175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr
            180             185             190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
```

<pre>
        195   200   205
    Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
        210   215   220

    Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
        225   230   235   240

    Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
                    245   250   255

    Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
                260   265   270

    Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
            275   280   285

    Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
        290   295   300
</pre>

(2) INFORMATION FOR SEQ ID NO:28:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

<pre>
    Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
    1   5   10   15

    Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
                20   25   30

    Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
            35   40   45

    Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
        50   55   60

    Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
    65   70   75   80

    Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Ser Phe Ser Gly Thr Val
                85   90   95

    Lys Tyr Asn Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
                100   105   110

    Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
            115   120   125
</pre>

```
Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130             135             140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145             150             155             160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165             170             175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala Cys Asp Val Thr Val Thr
            180             185             190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
            195             200             205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210             215             220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225             230             235             240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245             250             255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260             265             270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
    275             280             285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290             295             300
```

(2) INFORMATION FOR SEQ ID NO:29:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5               10              15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20              25              30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
            35              40              45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50              55              60
```

```
Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65              70          75                  80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Ser Phe Ser Glu Thr Val
            85              90                  95

Lys Tyr Asn Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100             105             110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115             120             125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130             135             140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145             150             155             160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165             170             175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr
        180             185             190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
        195             200             205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr Asp Ala Asp
    210             215             220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225             230             235             240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245             250             255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260             265             270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
    275             280             285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290             295             300
```

(2) INFORMATION FOR SEQ ID NO:30:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 300 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: unknown
       (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Lys Arg Val Ile Asn Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1   5   10   15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20   25   30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
        35   40   45

Val Asn Val Gly Gln His Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50   55   60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65   70   75   80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
            85   90   95

Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Leu
        100   105   110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115   120   125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130   135   140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145   150   155   160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165   170   175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr
            180   185   190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
        195   200   205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210   215   220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225   230   235   240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245   250   255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260   265   270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
    275   280   285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290   295   300

(2) INFORMATION FOR SEQ ID NO:31:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 296 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: peptide

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5                   10                  15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20              25                  30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
        35              40                  45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50              55                  60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65              70                  75                  80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
            85                  90                  95

Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100             105                 110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115             120                 125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Lys Ala Gly Ser Leu Ile Ala
    130             135                 140

Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser Asp Asp Phe Gln
145             150                 155                 160

Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val Val Pro Thr Gly
            165                 170                 175

Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr Leu Pro Asp Tyr
        180             185                 190

Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys Ala Lys Ser Gln
        195             200                 205

Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp Ala Gly Asn Ser
    210             215                 220

Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln Gly Val Gly Val
225             230                 235                 240

Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn Asn Thr Val Ser
```

```
                     245                      250                      255

     Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly Leu Thr Ala Asn
                 260                      265                      270

     Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn Val Gln Ser Ile
                 275                      280                      285

     Ile Gly Val Thr Phe Val Tyr Gln
                 290                      295
```

(2) INFORMATION FOR SEQ ID NO:32:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
     Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
     1               5                   10                  15

     Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
                 20                  25                  30

     Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Val
             35                  40                  45

     Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
         50                  55                  60

     Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Arg Gln
     65                  70                  75                  80

     Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
                 85                  90                  95

     Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
                 100                 105                 110

     Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
                 115                 120                 125

     Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
         130                 135                 140

     Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
     145                 150                 155                 160

     Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
                 165                 170                 175
```

```
Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr
            180                 185                 190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
        195                 200                 205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210                 215                 220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225                 230                 235                 240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
                245                 250                 255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260                 265                 270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
        275                 280                 285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290                 295                 300
```

(2) INFORMATION FOR SEQ ID NO:33:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
Met Lys Arg Val Ile Asn Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5               10              15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20              25              30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
        35              40              45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50              55              60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65              70              75              80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
            85              90              95

Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100             105             110
```

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115   120   125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Val Ile Lys Ala Gly
    130   135   140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145   150   155   160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
                165   170   175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr
            180   185   190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
        195   200   205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210   215   220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225   230   235   240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245   250   255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260   265   270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
        275   280   285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290   295   300

(2) INFORMATION FOR SEQ ID NO:34:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1   5   10   15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
        20   25   30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
        35   40   45

```
Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50              55              60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65              70              75                  80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Ser Phe Ser Gly Thr Val
                85              90                  95

Lys Tyr Asn Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100             105             110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115             120             125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130             135             140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145             150             155             160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165             170             175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr
            180             185             190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
        195             200             205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210             215             220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225             230             235             240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245             250             255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260             265             270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
        275             280             285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290             295             300
```

(2) INFORMATION FOR SEQ ID NO:35:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```
Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5                   10                  15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20              25                  30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
        35              40                  45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50              55                  60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65              70                  75                      80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
            85                  90                  95

Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
        100             105                 110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115                 120                 125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130                 135                 140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145                 150                 155                 160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165                 170                 175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala His Asp Val Thr Val Thr
            180                 185                 190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
        195                 200                 205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210                 215                 220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225                 230                 235                 240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245                 250                 255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260                 265                 270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
        275                 280                 285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
```

290　　　　　　　295　　　　　　300

(2) INFORMATION FOR SEQ ID NO:36:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 300 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: unknown
  (D) TOPOLOGY: linear

 (ii) MOLECULE TYPE: peptide

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1　　　　　　5　　　　　　　　10　　　　　　15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
　　　　20　　　　　　　25　　　　　　30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
　　　　35　　　　　　40　　　　　　45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
　　50　　　　　　55　　　　　　60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65　　　　　　70　　　　　　75　　　　　　80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
　　　　85　　　　　　90　　　　　　95

Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
　　　　100　　　　　　105　　　　　　110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Arg
　　　　115　　　　　　120　　　　　　125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
　　　　130　　　　　　135　　　　　　140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145　　　　　　150　　　　　　155　　　　　　160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
　　　　165　　　　　　170　　　　　　175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala Arg Asp Val Thr Val Thr
　　　　180　　　　　　185　　　　　　190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
　　　　195　　　　　　200　　　　　　205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
　　　　210　　　　　　215　　　　　　220

```
Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225             230             235             240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245             250             255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260             265             270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
            275             280             285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290             295             300
```

(2) INFORMATION FOR SEQ ID NO:37:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```
Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5               10              15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20              25              30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
            35              40              45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50              55              60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65              70              75              80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
            85              90              95

Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100             105             110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
            115             120             125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130             135             140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145             150             155             160
```

```
Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165             170                 175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala His Asp Val Thr Val Thr
            180             185                 190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
            195             200             205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210                 215             220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225                 230             235                 240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245             250                 255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260             265                 270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
            275             280             285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290             295             300
```

(2) INFORMATION FOR SEQ ID NO:38:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```
Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5                   10                  15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20                  25                  30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
            35                  40                  45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50                  55                  60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65                  70                  75                  80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Asn Phe Ser Gly Thr Val
            85                  90                  95
```

```
Lys Tyr Ser Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100             105             110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
            115             120             125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130             135             140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145             150             155             160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165             170             175

Val Pro Thr Gly Gly Cys Asp Val Ser Ala His Asp Val Thr Val Thr
            180             185             190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
            195             200             205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210             215             220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225             230             235             240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245             250             255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260             265             270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
            275             280             285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290             295             300
```

(2) INFORMATION FOR SEQ ID NO:39:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```
Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5               10              15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
            20              25              30
```

```
Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
        35              40                  45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
    50              55                  60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65              70                  75                      80

Arg Gly Ser Ala Tyr Gly Gly Val Leu Ser Ser Phe Ser Gly Thr Val
                85              90                  95

Lys Tyr Asn Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100             105             110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115             120             125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130             135                 140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145             150                 155                     160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
            165             170                     175

Val Pro Thr Gly Gly Cys Asp Ala Ser Ala Arg Asp Val Thr Val Thr
            180             185                 190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
        195             200             205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210             215             220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225             230                 235                 240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245             250             255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260             265             270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
        275             280             285

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
    290             295             300
```

(2) INFORMATION FOR SEQ ID NO:40:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```
Met Lys Arg Val Ile Thr Leu Phe Ala Val Leu Leu Met Gly Trp Ser
1               5                   10                  15

Val Asn Ala Trp Ser Phe Ala Cys Lys Thr Ala Asn Gly Thr Ala Ile
                20                  25                  30

Pro Ile Gly Gly Gly Ser Ala Asn Val Tyr Val Asn Leu Ala Pro Ala
            35                  40                  45

Val Asn Val Gly Gln Asn Leu Val Val Asp Leu Ser Thr Gln Ile Phe
        50                  55                  60

Cys His Asn Asp Tyr Pro Glu Thr Ile Thr Asp Tyr Val Thr Leu Gln
65                  70                  75                  80

Arg Gly Ser Ala Tyr Gly Asp Val Leu Ser Ser Phe Ser Gly Thr Val
                85                  90                  95

Lys Tyr Asn Gly Ser Ser Tyr Pro Phe Pro Thr Thr Ser Glu Thr Pro
            100                 105                 110

Arg Val Val Tyr Asn Ser Arg Thr Asp Lys Pro Trp Pro Val Ala Leu
        115                 120                 125

Tyr Leu Thr Pro Val Ser Ser Ala Gly Gly Val Ala Ile Lys Ala Gly
    130                 135                 140

Ser Leu Ile Ala Val Leu Ile Leu Arg Gln Thr Asn Asn Tyr Asn Ser
145                 150                 155                 160

Asp Asp Phe Gln Phe Val Trp Asn Ile Tyr Ala Asn Asn Asp Val Val
                165                 170                 175

Val Pro Thr Gly Gly Cys Asp Ala Ser Ala Arg Asp Val Thr Val Thr
            180                 185                 190

Leu Pro Asp Tyr Pro Gly Ser Val Pro Ile Pro Leu Thr Val Tyr Cys
            195                 200                 205

Ala Lys Ser Gln Asn Leu Gly Tyr Tyr Leu Ser Gly Thr His Ala Asp
    210                 215                 220

Ala Gly Asn Ser Ile Phe Thr Asn Thr Ala Ser Phe Ser Pro Ala Gln
225                 230                 235                 240

Gly Val Gly Val Gln Leu Thr Arg Asn Gly Thr Ile Ile Pro Ala Asn
            245                 250                 255

Asn Thr Val Ser Leu Gly Ala Val Gly Thr Ser Ala Val Ser Leu Gly
            260                 265                 270

Leu Thr Ala Asn Tyr Ala Arg Thr Gly Gly Gln Val Thr Ala Gly Asn
```

<pre>   275   280   285
</pre>

Val Gln Ser Ile Ile Gly Val Thr Phe Val Tyr Gln
<pre>   290   295   300
</pre>

(2) INFORMATION FOR SEQ ID NO:41:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 900 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```
ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG      60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCTATCCCTA TTGGCGGTGG CAGCGCCAAT     120

GTTTATGTAA ACCTTGCGCC CGTCGTGAAT GTGGGGCAAA ACCTGGTCGT GGATCTTTCG     180

ACGCAAATCT TTTGCCATAA CGATTATCCG GAAACCATTA CAGACTATGT CACACTGCAA     240

CGAGGCTCGG CTTATGGCGG CGTGTTATCT AATTTTTCCG GGACCGTAAA ATATAGTGGC     300

AGTAGCTATC CATTTCCTAC CACCAGCGAA ACGCCGCGCG TTGTTTATAA TTCGAGAACG     360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG CGGGGTGGCG     420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC     480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCTACTGGC     540

GGCTGCGATG TTTCTGCTCG TGATGTCACC GTTACTCTGC CGGACTACCG TGGTTCAGTG     600

CCAATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC     660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCTGCACAG     720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG     780

TTAGGAGCAG TAGGGACTTC GGCGGTGAGT CTGGGATTAA CGGCAAATTA TGCACGTACC     840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA     900
```

(2) INFORMATION FOR SEQ ID NO:42:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 900 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```
ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG        60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCAATCCCTA TTGGCGGTGG CAGCGCCAAT       120

GTTTATGTAA ACCTTGCGCC TGCCGTGAAT GTGGGGCAAA ACCTGGTCGT AGATCTTTCG       180

ACGCAAATCT TTTGCCATAA CGATTACCCA GAAACCATTA CAGACTATGT CACACTGCAA       240

CGAGGTTCGG CTTATGGCGG CGTGTTATCT AGTTTTTCCG GGACCGTAAA ATATAATGGC       300

AGTAGCTATC CTTTCCCTAC TACCAGCGAA ACGCCGCGGG TTGTTTATAA TTCGAGAACG       360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCGGTGA GCAGTGCGGG GGGAGTGGCG       420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC       480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCCACTGGC       540

GGCTGCGATG TTTCTGCTCG TGATGTCACC GTTACTCTGC CGGACTACCC TGGTTCAGTG       600

CCGATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC       660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCCGCGCAG       720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG       780

TTAGGAGCAG TAGGGACTTC GGCGGTAAGT CTGGGATTAA CGGCAAATTA CGCACGTACC       840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA       900
```


(2) INFORMATION FOR SEQ ID NO:43:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 900 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```
ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG        60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCAATCCCTA TTGGCGGTGG CAGCGCCAAT       120

GTTTATGTAA ACCTTGCGCC TGCCGTGAAT GTGGGGCAAA ACCTGGTCGT AGATCTTTCG       180

ACGCAAATCT TTTGCCATAA CGATTACCCA GAAACCATTA CAGACTATGT CACACTGCAA       240
```

CGAGGTTCGG CTTATGGCGG CGTGTTATCT AGTTTTTCCG GGACCGTAAA ATATAATGGC   300

AGTAGCTATC CTTTCCCTAC TACCAGCGAA ACGCCGCGGG TTGTTTATAA TTCGAGAACG   360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCGGTGA GCAGTGCGGG GGGAGTGGCG   420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC   480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCCACTGGC   540

GGCTGTGATG CTTCTGCTCG TGATGTCACC GTTACTCTGC CGGACTACCC TGGTTCAGTG   600

CCGATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTATCCGGC   660

ACACATGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCCGCGCAG   720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG   780

TTAGGAGCAG TAGGGACTTC GGCGGTGAGT CTGGGATTAA CGGCAAATTA TGCACGTACC   840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA   900

(2) INFORMATION FOR SEQ ID NO:44:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 900 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG   60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCTATCCCTA TTGGCGGTGG CAGCGCCAAT   120

GTTTATGTAA ACCTTGCGCC TGCCGTGAAT GTGGGGCAAA ACCTGGTCGT GGATCTTTCG   180

ACGCAAATCT TTTGCCATAA CGATTACCCG GAAACCATTA CAGACTATGT CACACTGCAA   240

CGAGGTTCGG CTTATGGCGG CGTGTTATCT AGTTTTTCCG AGACCGTAAA ATATAATGGC   300

AGTAGCTATC CTTTCCCTAC TACCAGCGAA ACGCCGCGGG TTGTTTATAA TTCGAGAACG   360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG GGGAGTGGCG   420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC   480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCCACTGGC   540

GGCTGTGATG TTTCTGCTCG TGATGTCACC GTTACTTTGC CGGACTACCC TGGTTCAGTG   600

CCGATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC   660

ACAGACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCTGCACAG   720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG   780

TTAGGAGCAG TAGGGACTTC GGCGGTAAGT CTGGGATTAA CGGCAAATTA CGCACGTACC   840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA   900

(2) INFORMATION FOR SEQ ID NO:45:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 900 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG CTGGTCGGT AAATGCCTGG   60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCAATCCCTA TTGGCGGTGG CAGCGCCAAT   120

GTTTATGTAA ACCTTGCGCC TGCCGTGAAT GTGGGGCAAA ACCTGGTCGT AGATCTTTCG   180

ACGCAAATCT TTTGCCATAA CGATTACCCA GAAACCATTA CAGACTATGT CACACTGCAA   240

CGAGGTTCGG CTTATGGCGA CGTGTTATCT AGTTTTTCCG GGACCGTAAA ATATAATGGC   300

AGTAGCTATC CTTTCCCTAC TACCAGCGAA ACGCCGCGGG TTGTTTATAA TTCGAGAACG   360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCGGTGA GCAGTGCGGG GGGAGTGGCG   420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC   480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCCACTGGC   540

GGCTGTGATG TCTCTGCTCG TGATGTCACC GTTACTCTGC CGGACTACCC TGGTTCAGTG   600

CCGATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTATCCGGC   660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCCGCGCAG   720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG   780

TTAGGAGCAG TAGGGACTTC GGCGGTAAGT CTGGGATTAA CGGCAAATTA CGCACGTACC   840

GGAGGGCAGG TGACCGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA   900

(2) INFORMATION FOR SEQ ID NO:46:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 900 base pairs

      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: DNA (genomic)

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG       60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCTATTCCTA TTGGCGGTGG CAGCGCTAAT      120

GTTTATGTAA ACCTTGCGCC TGCCGTGAAT GTGGGGCAAA ACCTGGTCGT AGATCTTTCG      180

ACGCAAATCT TTTGCCATAA CGATTATCCG GAAACCATTA CAGACTATGT CACACTGCAA      240

CGAGGCTCGG CTTATGGCGG CGTGTTATCT AATTTTTCCG GGACCGTAAA ATATAGTGGC      300

AGTAGCTATC CATTCCCGAC TACCAGCGAA ACGCCGCGGG TTGTTTATAA TTCGAGAACG      360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG TGGGGTGGCG      420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC      480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCTACTGGC      540

GGCTGCGATG TTTCTGCTCA TGATGTCACC GTTACTCTGC CGGACTACCC TGGTTCAGTG      600

CCAATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC      660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCAGCGCAG      720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG      780

TTAGGAGCAG TAGGGACTTC GGCGGTAAGT CTGGGATTAA CGGCAAATTA CGCACGTACC      840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA      900
```

(2) INFORMATION FOR SEQ ID NO:47:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 900 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: DNA (genomic)

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```
ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG       60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCTATTCCTA TTGGCGGTGG CAGCGCTAAT      120
```

```
GTTTATGTAA ACCTTGCGCC TGCCGTGAAT GTGGGGCAAA ACCTGGTCGT AGATCTTTCG      180

ACGCAAATCT TTTGCCATAA CGATTATCCG GAAACCATTA CAGACTATGT CACACTGCAA      240

CGAGGCTCGG CTTATGGCGG CGTGTTATCT AATTTTTCCG GGACCGTAAA ATATAGTGGC      300

AGTAGCTATC CATTTCCGAC TACCAGCGAA ACGCCGCGGG TTGTTTATAA TTCGAGAACG      360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG TGGGGTGGCG      420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC      480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCTACTGGC      540

GGCTGCGATG TTTCTGCTCA TGATGTCACC GTTACTCTGC CGGACTACCC TGGTTCAGTG      600

CCAATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC      660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCAGCGCAG      720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG      780

TTAGGAGCAG TAGGGACTTC GGCGGTAAGT CTGGGATTAA CGGCAAATTA CGCACGTACC      840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA      900
```

(2) INFORMATION FOR SEQ ID NO:48:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 888 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

```
ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG       60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCTATCCCTA TTGGCGGTGG CAGCGCCAAT      120

GTTTATGTAA ACCTTGCGCC CGCCGTGAAT GTGGGGCAAA ACCTGGTCGT GGATCTTTCG      180

ACGCAAATCT TTTGCCATAA CGATTATCCG GAAACCATTA CAGACTATGT CACACTGCAA      240

CGAGGCTCGG CTTATGGCGG CGTGTTATCT AATTTTTCCG GGACCGTAAA ATATAGTGGC      300

AGTAGCTATC CATTTCCTAC CACCAGCGAA ACGCCGCGCG TTGTTTATAA TTCGAGAACG      360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG TAAAGCTGGC      420

TCATTAATTG CCGTGCTTAT TTTGCGACAG ACCAACAACT ATAACAGCGA TGATTTCCAG      480

TTTGTGTGGA ATATTTACGC CAATAATGAT GTGGTGGTGC CTACTGGCGG CTGCGATGTT      540
```

TCTGCTCGGG ATGTCACCGT TACTCTGCCG GACTACCCTG GTTCAGTGCC AATTCCTCTT   600

ACCGTTTATT GTGCGAAAAG CCAAAACCTG GGGTATTACC TCTCCGGCAC ACACGCAGAT   660

GCGGGCAACT CGATTTTCAC CAATACCGCG TCGTTTTCAC CTGCACAGGG CGTCGGCGTA   720

CAGTTGACGC GCAACGGTAC GATTATTCCA GCGAATAACA CGGTATCGTT AGGAGCAGTA   780

GGGACTTCGG CGGTGAGTTT GGGATTAACG GCAAATTATG CACGTACCGG AGGGCAGGTG   840

ACTGCAGGGA ATGTGCAATC GATTATTGGC GTGACTTTTG TTTATCAA   888

(2) INFORMATION FOR SEQ ID NO:49:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 900 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG   60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCTATCCCTA TTGGCGGTGG CAGCGCCAAT   120

GTTTATGTAA ACCTTGCGCC CGTCGTGAAT GTGGGGCAAA ACCTGGTCGT GGATCTTTCG   180

ACGCAAATCT TTTGCCATAA CGATTATCCG GAAACCATTA CAGACTATGT CACACTGCAA   240

CGAGGCTCGG CTTATGGCGG CGTGTTATCT AATTTTTCCG GGACCGTAAA ATATAGTGGC   300

AGTAGCTATC CATTTCCTAC CACCAGCGAA ACGCCGCGCG TTGTTTATAA TTCGAGAACG   360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG CGGGGTGGCG   420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC   480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCTACTGGC   540

GGCTGCGATG TTTCTGCTCG TGATGTCACC GTTACTCTGC CGGACTACCC TGGTTCAGTG   600

CCAATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC   660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCTGCACAG   720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGAATATTC CAGCGAATAA CACGGTATCG   780

TTAGGAGCAG TAGGGACTTC GGCGGTGAGT CTGGGATTAA CGGCAAATTA TGCACGTACC   840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA   900

(2) INFORMATION FOR SEQ ID NO:50:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 900 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

ATGAAACGAG TTATTAACCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG   60

TCATTCGCCT GTAAAACCGC CAATGGCACC GCTATCCCTA TTGGCGGTGG CAGCGCCAAT   120

GTTTATGTAA ACCTTGCGCC CGCCGTGAAT GTGGGGCAAA ACCTGGTCGT GGATCTTTCG   180

ACGCAAATCT TTTGCCATAA CGATTACCCG GAAACCATTA CAGATTATGT CACACTGCAA   240

CGAGGCTCGG CTTATGGCGG CGTGTTATCT AATTTTTCCG GGACCGTAAA ATATAGTGGC   300

AGTAGCTATC CATTTCCGAC CACCAGTGAA ACGCCGCGGG TTGTTTATAA TTCGAGAACG   360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG CGGGGTGGTG   420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC   480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCCACTGGC   540

GGCTGCGATG TTTCTGCTCG TGATGTCACC GTTACTCTGC CGGACTACCC TGGTTCAGTG   600

CCGATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC   660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCTGCACAG   720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG   780

TTAGGAGCAG TAGGGACTTC GGCGGTAAGT CTGGGATTAA CGGCAAATTA CGCACGTACC   840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG CCGTGACTTT TGTTTATCAA   900

(2) INFORMATION FOR SEQ ID NO:51:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 900 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

ATGAAACGAG TTATTAACCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG   60

```
TCATTCGCCT GTAAAACCGC CAATGGCACC GCTATCCCTA TTGGCGGTGG CAGCGCCAAT      120

GTTTATGTAA ACCTTGCGCC CGCCGTGAAT GTGGGGCAAC ACCTGGTCGT AGATCTTTCG      180

ACGCAAATCT TTTGCCATAA CGATTACCCG GAAACCATTA CAGACTATGT CACACTGCAA      240

CGAGGTTCGG CTTATGGCGG CGTGTTATCT AATTTTTCCG GGACCGTAAA ATATAGTGGC      300

AGTAGCTATC CATTTCCTAC CACCAGCGAA ACGCTGCGGG TTGTTTATAA TTCGAGAACG      360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG CGGGGTGGCG      420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC      480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCTACTGGC      540

GGCTGCGATG TTTCTGCTCG TGATGTCACC GTTACTCTGC CGGACTACCC TGGTTCAGTG      600

CCAATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC      660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CCTCGTTTTC ACCAGCGCAG      720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG      780

TTAGGAGCAG TAGGGACTTC GGCGGTAAGT CTGGGATTAA CGGCAAATTA CGCACGTACC      840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA      900
```

(2) INFORMATION FOR SEQ ID NO:52:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 900 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```
ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG CTGGTCGGT AAATGCCTGG      60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCTATCCCTA TTGGCGGTGG CAGCGCTAAT      120

GTTTATGTAA ACCTTGCGCC TGCCGTGAAT GTGGGGCAAA ACCTGGTCGT AGATCTTTCG      180

ACGCAAATCT TTTGCCATAA CGATTATCCG GAAACCATTA CAGACTATGT CACACTGCAA      240

CGAGGCTCGG CTTATGGCGG CGTGTTATCT AATTTTTCCG GGACCGTAAA ATATAGTGGC      300

AGTAGCTATC CATTTCCGAC TACCAGCGAA ACGCCGCGGG TTGTTTATAA TTCGAGAACG      360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG TGGGGTGGCG      420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC      480
```

```
GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCTACTGGC      540

GGCTGCGATG TTTCTGCTCA TGATGTCACC GTTACTCTGC CGGACTACCC TGGTTCAGTG      600

CCAATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC      660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCAGCGCAG      720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG      780

TTAGGAGCAG TAGGGACTTC GGCGGTAAGT CTGGGATTAA CGGCAAATTA CGCACGTACC      840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA      900
```

(2) INFORMATION FOR SEQ ID NO:53:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 900 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

```
ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG CTGGTCGGT AAATGCCTGG       60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCTATCCCTA TTGGCGGTGG CAGCGCCAAT     120

GTTTATGTAA ACCTTGCGCC CGCCGTGAAT GTGGGGCAAA ACCTGGTCGT GGATCTTTCG     180

ACGCAAATCT TTTGCCATAA CGATTATCCG GAAACCATTA CAGACTATGT CACACTGCAA     240

CGAGGCTCGG CTTATGGCGG CGTGTTATCT AATTTTTCCG GGACCGTAAA ATATAGTGGC     300

AGTAGCTATC CATTTCCTAC CACCAGCGAA ACGCCGCGCG TTGTTTATAA TTCGAGAACG     360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG CGGGGTGGCG     420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC     480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCTACTGGC     540

GGCTGCGATG TTTCTGCTCG TGATGTCACC GTTACTCTGC CGGACTACCC AGGTTCAGTG     600

CCAATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC     660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCTGCACAG     720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG     780

TTAGGAGCAG TAGGGACTTC GGCGGTGAGT CTGGGATTAA CGGCAAATTA TGCACGTACC     840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA     900
```

(2) INFORMATION FOR SEQ ID NO:54:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 900 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```
ATGAAACGAG TTATTACCCT GTTTGCTGTA CTGCTGATGG GCTGGTCGGT AAATGCCTGG      60

TCATTCGCCT GTAAAACCGC CAATGGTACC GCTATCCCTA TTGGCGGTGG CAGCGCCAAT     120

GTTTATGTAA ACCTTGCGCC TGCCGTGAAT GTGGGGCAAA ACCTGGTCGT GGATCTTTCG     180

ACGCAAATCT TTTGCCATAA CGATTACCCG GAAACCATTA CAGACTATGT CACACTGCAA     240

CGAGGTTCGG CTTATGGCGG CGTGTTATCT AGTTTTTCCG GGACCGTAAA ATATAATGGC     300

AGTAGCTATC CTTTCCCTAC TACCAGCGAA ACGCCGCGCG TTGTTTATAA TTCGAGAACG     360

GATAAGCCGT GGCCGGTGGC GCTTTATTTG ACGCCTGTGA GCAGTGCGGG GGGAGTGGCG     420

ATTAAAGCTG GCTCATTAAT TGCCGTGCTT ATTTTGCGAC AGACCAACAA CTATAACAGC     480

GATGATTTCC AGTTTGTGTG GAATATTTAC GCCAATAATG ATGTGGTGGT GCCCACTGGC     540

GGCTGTGATG TTTCTGCTTG TGATGTCACC GTTACTTTGC CGGACTACCC TGGTTCAGTG     600

CCGATTCCTC TTACCGTTTA TTGTGCGAAA AGCCAAAACC TGGGGTATTA CCTCTCCGGC     660

ACACACGCAG ATGCGGGCAA CTCGATTTTC ACCAATACCG CGTCGTTTTC ACCTGCACAG     720

GGCGTCGGCG TACAGTTGAC GCGCAACGGT ACGATTATTC CAGCGAATAA CACGGTATCG     780

TTAGGAGCAG TAGGGACTTC GGCGGTAAGT CTGGGATTAA CGGCAAATTA CGCACGTACC     840

GGAGGGCAGG TGACTGCAGG GAATGTGCAA TCGATTATTG GCGTGACTTT TGTTTATCAA     900
```

(2) INFORMATION FOR SEQ ID NO:55:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 51 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GATCTGTTGA AGTTCCGGGT AGTCAGCATA TCGATAGTCA GAAAAAAGCT G   51

CLAIMS

1. A method of targeting a bacterial adhesin to a specific location, comprising (i) identifying in said location adhes-in-interacting receptor moiety which is recognizable by
5   bacterial adhesins, (ii) isolating a bacterial cell that grows in said location and expresses an adhesin recognizing and interacting with said receptor moiety, and administering to the location the bacterial cell or the adhesin under conditions where the adhesin and the receptor moiety are
10   brought into interacting contact whereby the adhesin is associated with the receptor moiety.

2. A method according to claim 1 wherein the receptor moiety is selected from the group consisting of a glycolipid, a glycoprotein, a protein, a polypeptide, a saccharide moiety
15   and a peptide.

3. A method according to claim 1 wherein the isolated bacte-rial cell expresses an adhesin having modified receptor moiety-binding properties relative to an adhesin natively expressed by the cell, the isolation of the cell comprising
20   identifying in a parent bacterial cell, DNA sequence(s) coding for the binding domain(s) of said natively expressed adhesin and substituting at least one codon herein, whereby a modified adhesin molecule is expressed that is different in at least one amino acid from the adhesin expressed natively,
25   and selecting a the bacterial cell expressing the modified adhesin having an altered adhesion phenotype relative to the natively expressed bacterial adhesin.

4. A method according to claim 1 wherein the bacterial cell expressing an adhesin that recognizes and binds to the
30   receptor moiety is a recombinant bacterial cell derived from a parent bacterial cell that does not produce an adhesin binding to said receptor, by inserting into the parent cell a DNA sequence coding for an adhesin binding to the receptor

moiety, and selecting a bacterial cell expressing the DNA
sequence.

5. A method according to claim 1 wherein a non-adhesin com-
pound is associated with the adhesin, whereby said compound
5   is targeted with the adhesin to the location comprising the
receptor moiety recognizable by the adhesin.

6. A method according to claim 5 wherein the compound is
covalently bound to the adhesin.

7. A method according to claim 6 wherein the compound is part
10   of a fusion protein comprising the adhesin, the compound
being selected from the group consisting of an enzyme, an
antibody, an epitope and a toxin.

8. A method according to claim 5 wherein the compound is one
associated with the adhesin by a non-covalent binding.

15   9. A method according to claim 8 wherein the compound is
selected from the group consisting of a pharmacologically
active, a diagnostically active and an imaging compound.

10. A method according to claim 1 wherein the specific loca-
tion is a human or animal surface.

20   11. A method according to claim 1 wherein the specific loca-
tion is a plant surface.

12. A method according to claim 1 wherein the bacterial cell
expresses a recombinant bacterial adhesin variant derived
from a naturally occurring parent adhesin, said recombinant
25   bacterial adhesin variant having altered binding properties
relative to the naturally occurring adhesin from which it is
derived, the altered binding properties including binding to
at least one receptor moiety to which the parent adhesin does
not bind.

13. A method according to claim 12 wherein the adhesin vari-
ant is derived from a naturally occurring adhesin isolated
from a cell structure selected from the group consisting of a
capsule, a lipopolysaccharide layer, an outer membrane pro-
5 tein, a flagellum, a pilus, a fimbria, a non-fimbrial adhesin
(NFA) and an afimbrial adhesin (AFA).

14. A method according to claim 12 or 13 wherein the adhesin
variant is a protein having an amino acid sequence differing
in at least one amino acid residue from its parent protein
10 adhesin.

15. A method according to claim 14 wherein the adhesin vari-
ant is a FimH adhesin having an amino acid sequence which
differs from the *E. coli* PC31 FimH adhesin as defined in
Table 1 herein in at least one amino acid.

15 16. A method according to claim 15 wherein the FimH adhesin
is one binding to a receptor selected from the group consist-
ing of a domain where mannosyl residues are not terminal and
a domain devoid of saccharide.

17. A method according to claim 15 wherein the adhesin vari-
20 ant is a chimeric adhesin comprising amino acid sequences
from different FimH adhesins.

18. A method according to claim 15 wherein the FimH adhesin
has an amino acid sequence which is selected from the group
consisting of sequences appearing in Fig. 5 herein with
25 designations CI#12, CI#4, CI#7 or CSH-50.

19. A method according to claim 15 wherein the adhesin is one
which, when tested for binding to yeast mannan (Mn), human
plasma fibronectin (Fn), periodate treated Fn and the syn-
thetic peptide FnSp1 comprising the first 30 amino acids of
30 Fn, only binds to Mn (M class).

20. A method according to claim 15 wherein the adhesin is one which, when tested for binding to yeast mannan (Mn), human plasma fibronectin (Fn), periodate treated Fn and the synthetic peptide FnSp1 comprising the first 30 amino acids of
5   Fn, binds to Mn and Fn (MF class).

21. A method according to claim 15 wherein the adhesin is one which, when tested for binding to yeast mannan (Mn), human plasma fibronectin (Fn), periodate treated Fn and the synthetic peptide FnSp1 comprising the first 30 amino acid
10 residues of Fn, binds to all of these (MFP class).

22. A method according to claim 15 wherein the adhesin is one which, when tested for binding to five Fn-fragments obtained by thermolysin treatment, only binds to the 40-kDa gelatin-binding fragment.

15 23. A method according to claim 22 wherein the adhesin is one which, when tested for binding to five Fn-fragments obtained by thermolysin treatment, binds to all five fragments.

24. A method according to claim 15 wherein the adhesin is at least 90% homologous to the PC31 FimH adhesin.

20 25. A method according to claim 15 wherein the adhesin is a chimeric adhesin comprising amino acid sequences from different FimH adhesins.

26. A method according to claim 15 comprising an amino acid sequence which differs from the *E. coli* PC31 FimH adhesin by
25 at least one amino acid occurring between residues 27 and 119 of the mature FimH sequence.

27. A method according to claim 15 wherein the adhesin binds to a receptor moiety selected from the group consisting of a human receptor moiety, an animal receptor moiety, a plant
30 receptor moiety and an inanimate, non-biological receptor moiety.

28. A method according to claim 1 wherein the bacterial cell being targeted is a cell comprising a gene coding for a gene product which, when expressed has a killing or cell function-limiting effect in said cell, the expression of said gene

5 coding for the cell killing or cell function-limiting gene product being regulated in such a manner that the bacterial cell when targeted will be killed or limited in its function in a pre-determined manner.

29. A recombinant or mutant bacterial adhesin variant derived

10 from a naturally occurring parent adhesin, said adhesin variant having altered binding properties relative to the naturally occurring adhesin from which it is derived, the altered binding properties including binding to at least one receptor to which the parent adhesin does not bind.

15 30. An adhesin variant according to claim 29 which is derived from a naturally occurring adhesin isolated from a cell structure selected from the group consisting of a capsule, a lipopolysaccharide layer, an outer membrane protein, a flagellum, a pilus, a fimbria, a non-fimbrial adhesin (NFA)

20 and an afimbrial adhesin (AFA).

31. An adhesin variant according to claim 29 or 30 which is a protein having an amino acid sequence differing by at least one amino acid residue from its parent protein adhesin.

32. An adhesin variant according to claim 29 which is a FimH

25 mannose-sensitive adhesin having an amino acid sequence which differs from the *E. coli* PC31 FimH adhesin as defined in Table 1 herein by at least one amino acid, said FimH adhesin binding to a receptor selected from the group consisting of a domain where mannosyl residues are not terminal and a domain

30 devoid of saccharide.

33. An adhesin variant according to claim 32 which is at least 90% homologous to the PC31 FimH adhesin.

34. An adhesin variant according to claim 32 which is a
chimeric adhesin comprising amino acid sequences from differ-
ent FimH adhesins.

35. An adhesin variant according to claim 29 which binds to a
5 receptor moiety selected from the group consisting of an
animal receptor moiety, a plant receptor moiety and an inani-
mate receptor moiety.

36. An adhesin variant according to claim 29 which is part of
a fusion protein comprising the adhesin variant and a
10 heterologous polypeptide.

37 An adhesin variant according to claim 36 wherein the
heterologous polypeptide is selected from the group consist-
ing of an epitope, an enzyme, a toxic gene product and an
antibody.

15 38. A FimH adhesin comprising 279 amino acids, having an
amino acid sequence which differs from the *E. coli* PC31 FimH
adhesin as defined in Table 1 herein by at least one amino
acid.

39. A FimH adhesin according to claim 38 which has an amino
20 acid sequence which is selected from the group of sequences
appearing in Fig. 5 herein with designations CI#12, CI#4,
CI#7 or CSH-50.

40. An adhesin according to claim 38 which binds to a
receptor selected from the group consisting of a domain where
25 mannosyl residues are not terminal, a domain devoid of
saccharide, a glycolipid, a glycoprotein, a protein, a
polypeptide and a peptide.

41. An adhesin according to claim 38 which when tested for
30 binding to yeast mannan (Mn), human plasma fibronectin (Fn),
periodate treated Fn and the synthetic peptide FnSp1 compris-

ing the first 30 amino acids of Fn, only binds to Mn (M class).

42. An adhesin according to claim 38 which when tested for binding to yeast mannan (Mn), human plasma fibronectin (Fn), periodate treated Fn and the synthetic peptide FnSp1 compris-
5 ing the first 30 amino acids of Fn, binds to Mn and Fn (MF class).

43. An adhesin according to claim 38 which when tested for binding to yeast mannan (Mn), human plasma fibronectin (Fn), periodate treated Fn and the synthetic peptide FnSp1 compris-
10 ing the first 30 amino acid residues of Fn, binds to all of these (MFP class).

44. An adhesin according to claim 38 which when tested for binding to five Fn-fragments obtained by thermolysin treatment, only binds to the 40-kDa gelatin-binding fragment.

15 45. An adhesin according to claim 38 which when tested for binding to five Fn-fragments obtained by thermolysin treatment, binds to all five fragments.

46. An adhesin according to claim 38 which is at least 90% homologous to the PC31 FimH adhesin.

20 47. An adhesin according to claim 38 which is a chimeric adhesin comprising amino acid sequences from different FimH adhesins.

48. An adhesin according to claim 38 comprising an amino acid sequence which differs from the *E. coli* PC31 FimH adhesin by
25 at least one amino acid occurring between residues 27 and 119 of the mature FimH sequence.

49. An adhesin according to claim 48 which binds to a receptor moiety selected from the group consisting of a human receptor moiety, an animal receptor moiety and a plant
30 receptor moiety.

50. A recombinant replicon comprising a DNA sequence selected
from the group consisting of a sequence coding for a recombi-
nant bacterial adhesin as defined in claim 29 and a sequence
coding for a FimH adhesin as defined in claim 38.

5   51. A recombinant replicon according to claim 50 wherein the
DNA sequence codes for a FimH adhesin having an amino acid
sequence which differs from the *E. coli* PC31 FimH adhesin by
at least one amino acid.

52. A replicon according to claim 52 in which the DNA
10   sequence is at least 90% homologous to the PC31 *fimH* gene.

53. A replicon according to claim 50 in which the DNA
sequence is a chimeric *fimH* gene comprising DNA from differ-
ent *fimH* genes.

54. A replicon according to claim 50 in which the DNA
15   sequence comprises a DNA sequence which differs from the *E.
coli* PC31 *fimH* gene by at least one codon between the codons
coding for amino acid residues 27 and 119 of the mature FimH
sequence.

55. A replicon according to claim 50 in which the DNA
20   sequence comprises a further DNA sequence coding for a
heterologous polypeptide.

56. A replicon according to claim 55 wherein the polypeptide
is selected from a group consisting of an epitope, an enzyme,
a toxic gene product and an antibody.

25   57. A replicon according to claim 55 wherein the further DNA
sequence codes for a gene product which is selected from a
pesticidally active gene product and a pollutant-degrading
gene product.

58. A replicon according to claim 50 wherein the DNA sequence
30   is isolated from an *Enterobacteriaceae* species.

59. A fusion protein comprising an adhesin selected from the group consisting of a recombinant bacterial adhesin variant as defined in claim 29 and a FimH adhesin as defined in claim 38, and a heterologous polypeptide.

5   60. A fusion protein according to claim 59 wherein the heterologous polypeptide is selected from an epitope, an enzyme, a toxic gene product and an antibody.

61. A fusion protein according to claim 59 which carries a non-covalently bound compound.

10   62. A bacterial cell which expresses an adhesin selected from the group consisting of a recombinant bacterial adhesin variant as defined in claim 29 and a FimH adhesin as defined in claim 38.

63. A recombinant bacterial cell according to claim 62 which 15 comprises a recombinant replicon as defined in claim 50.

64. A bacterial cell according to claim 62 which is selected from *Enterobacteriaceae*, *Pseudomonadaceae*, *Vibrionaceae* and *Baccilaceae*.

65. A bacterial cell according to claim 62 which further 20 expresses a gene product selected from the group consisting of a pesticidally active compound, an immunologically active gene product and a pollutant-degrading active compound.

66. A bacterial cell according to claim 62 in which the recombinant adhesin variant is expressed as a fusion protein 25 comprising the adhesin variant and a further polypeptide.

67. A bacterial cell according to claim 62 which further comprises a gene coding for a gene product which, when expressed has a killing or cell function-limiting effect in said cell, the expression of said gene coding for the cell

killing or cell function-limiting gene product being regu-
lated in such a manner that the bacterial cell when targeted
to receptor in a specific location will be killed or limited
in its function in a pre-determined manner.

5   68. A method of isolating a bacterial cell expressing an
     adhesin having modified binding properties relative to a
     natively expressed adhesin, comprising identifying in the
     bacterial cell DNA sequence(s) coding for the binding
     domain(s) of said natively expressed adhesin and substituting
10   at least one codon herein, whereby a modified adhesin mole-
     cule is expressed that is different in at least one amino
     acid from the adhesin expressed natively, and selecting a
     bacterial cell expressing the modified adhesin having an
     altered adhesion phenotype relative the natively expressed
15   bacterial adhesin.

     69. A method according to claim 68 wherein a non-adhesin
     compound is associated with the adhesin.

     70. A method according to claim 69 wherein the non-adhesin
     compound is associated with the adhesin by being expressed
20   with the adhesin as part of a fusion protein comprising the
     adhesin.

     71. A method according to claim 68 which in a further step
     comprises binding non-covalently a compound to the adhesin
     when expressed.

25   72. A method according to claim 68 wherein the natively
     expressed adhesin is a FimH adhesin.

     73. A method according to claim 68 wherein the codon(s)
     is/are substituted by mutagenization.

     74. A method of preparing a recombinant bacterial cell that
30   binds to a specific receptor moiety, comprising introducing
     into a bacterium that does not produce an adhesin binding to

said receptor moiety, a DNA sequence coding for an adhesin
binding to the receptor moiety, and selecting a bacterial
cell expressing the DNA sequence.

75. A method according to claim 74 wherein the DNA sequence
5   coding for an adhesin binding to the receptor moiety is a
sequence coding for a FimH adhesin.

76. A method according to claim 74 wherein the DNA is intro-
duced by transforming the bacterial cell with a recombinant
replicon as defined in claim 50.

10   77. A method according to claim 74 wherein a non-adhesin
compound is associated with the adhesin.

78. A method according to claim 77 wherein the non-adhesin
compound is associated with the adhesin by being expressed
with the adhesin as part of a fusion protein comprising the
15   adhesin.

79. A method according to claim 74 which in a further step
comprises binding non-covalently a compound to the adhesin
when expressed.

80. A method of providing a mutant bacterial cell having
20   fimbriae which binds to a moiety to which the wild-type cell
from which the mutant cell is derived does not bind, compris-
ing contacting a population of said wild-type cell with said
moiety, removing the contacted cells which do not bind to the
moiety, cultivating cells binding to the moiety to obtain a
25   culture which is enriched with regard to cells binding to the
moiety and selecting from said culture a mutant cell binding
to said moiety.

81. A method according to claim 80 wherein the moiety with
which the wild-type cell population is contacted, is a
30   casein.

82. A method of isolating a compound from a solution or
suspension containing the compound, the method comprising
contacting the solution or the suspension with a fusion
protein according to claim 59 wherein the heterologous
5   polypeptide has an affinity to the compound to be isolated.

83. A composition comprising a population of a bacterial cell
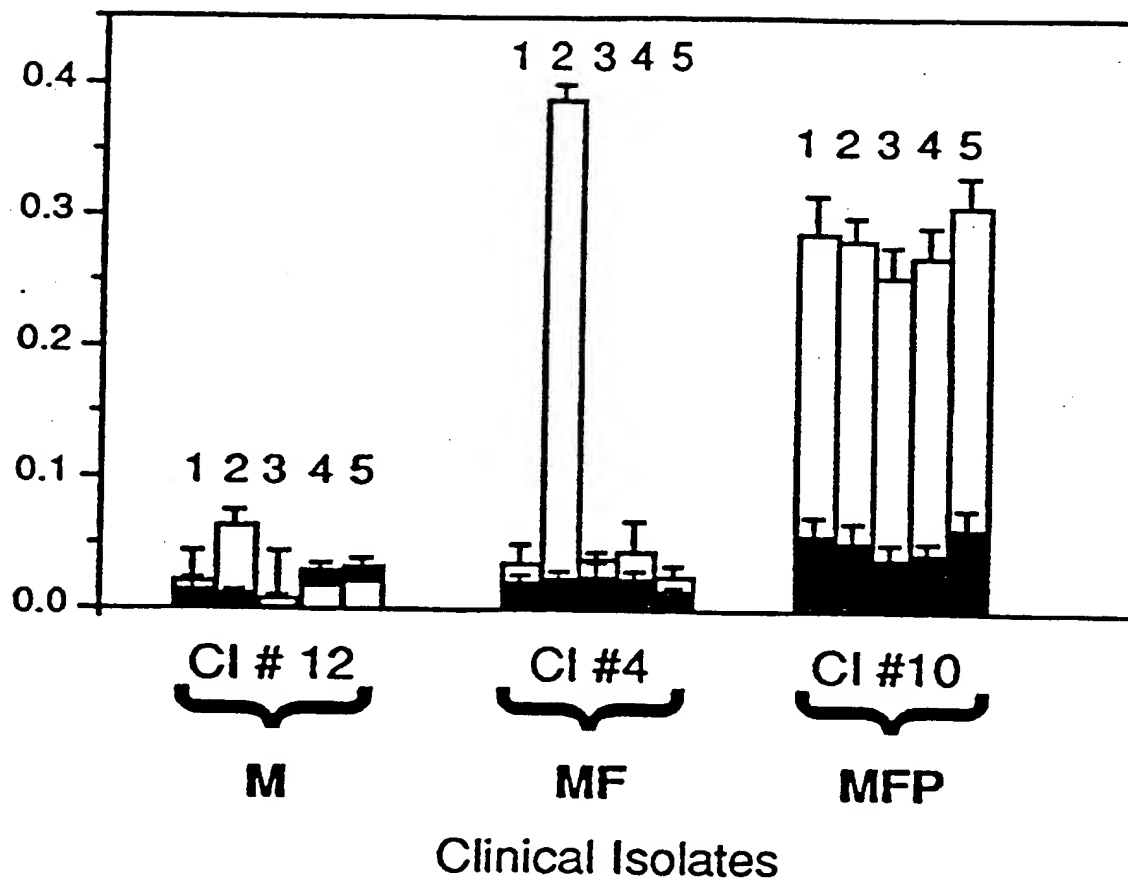as defined in claim 62.
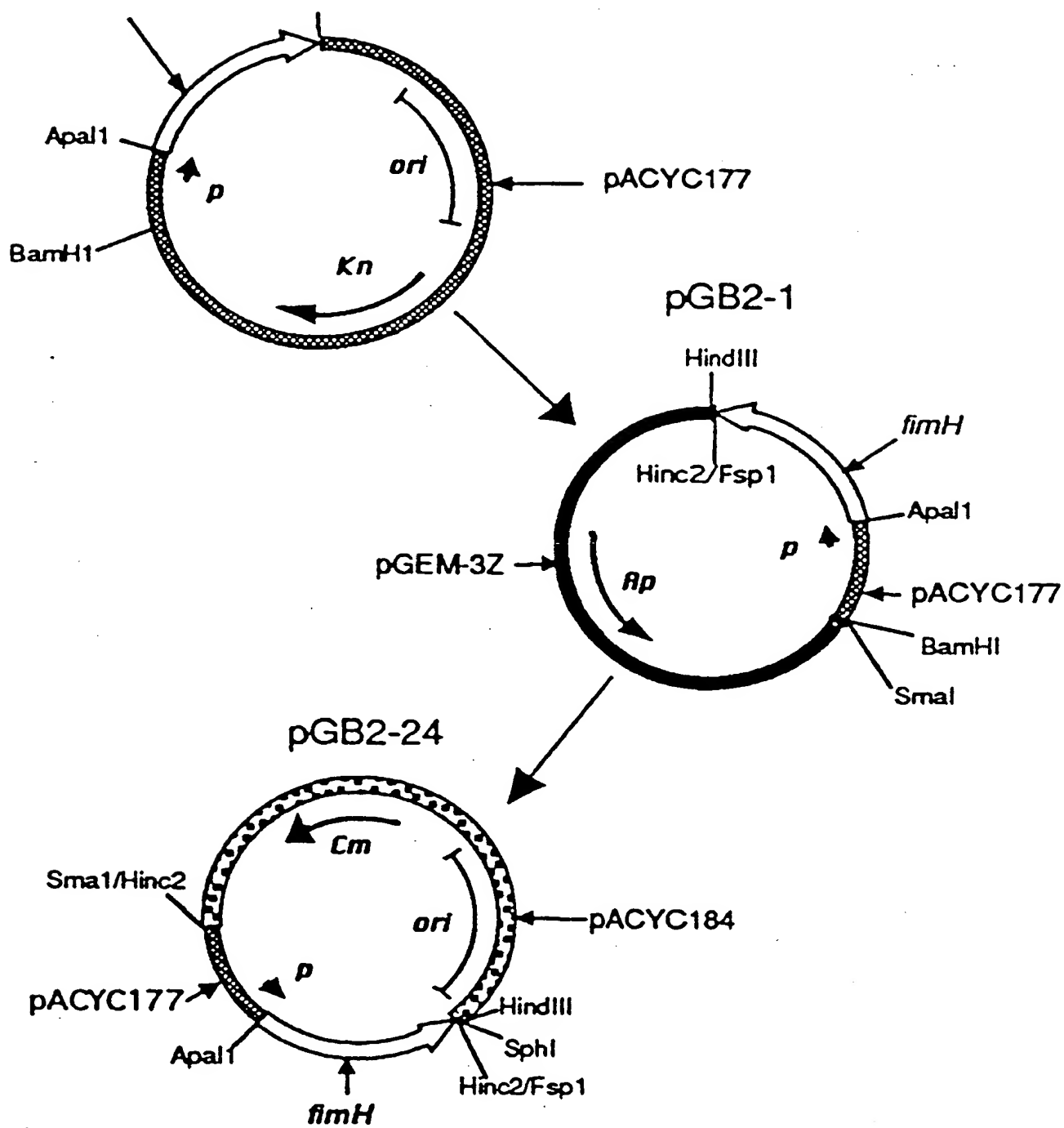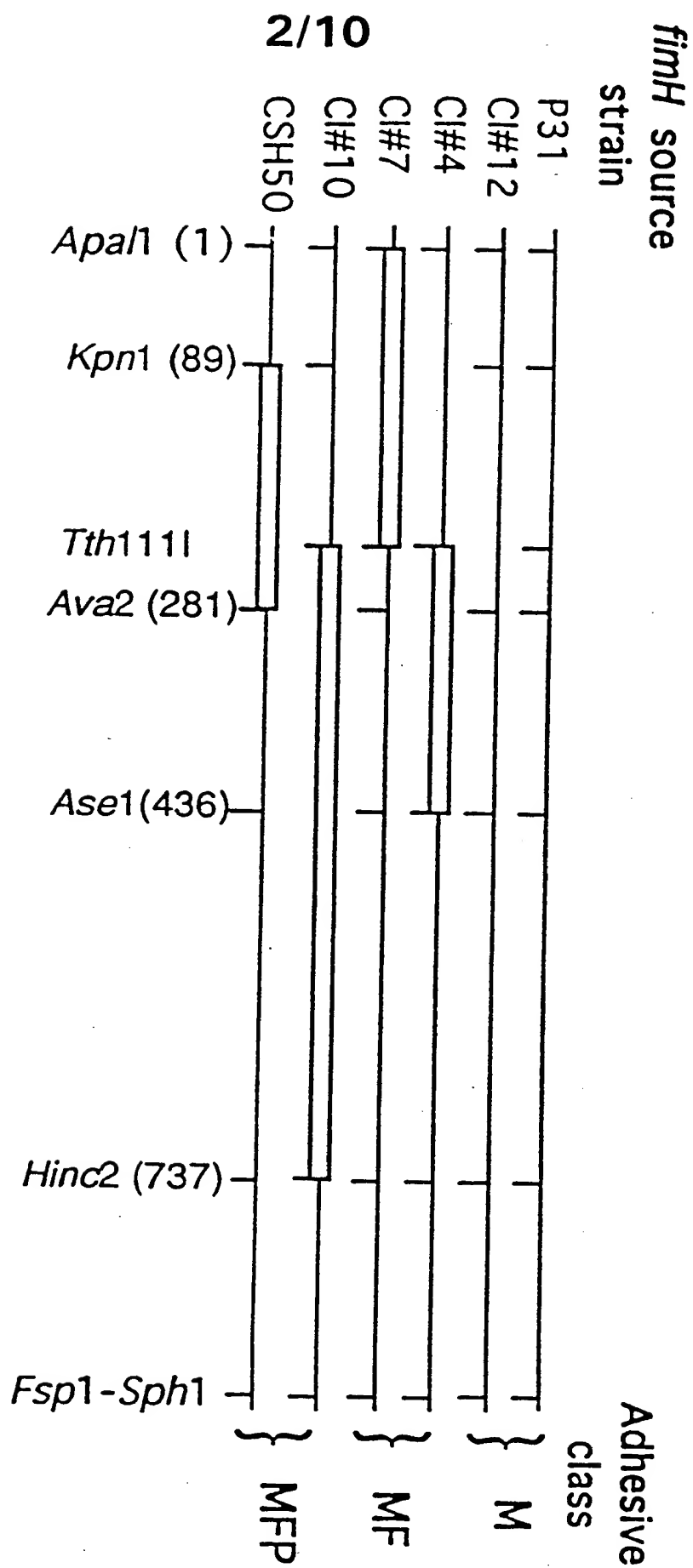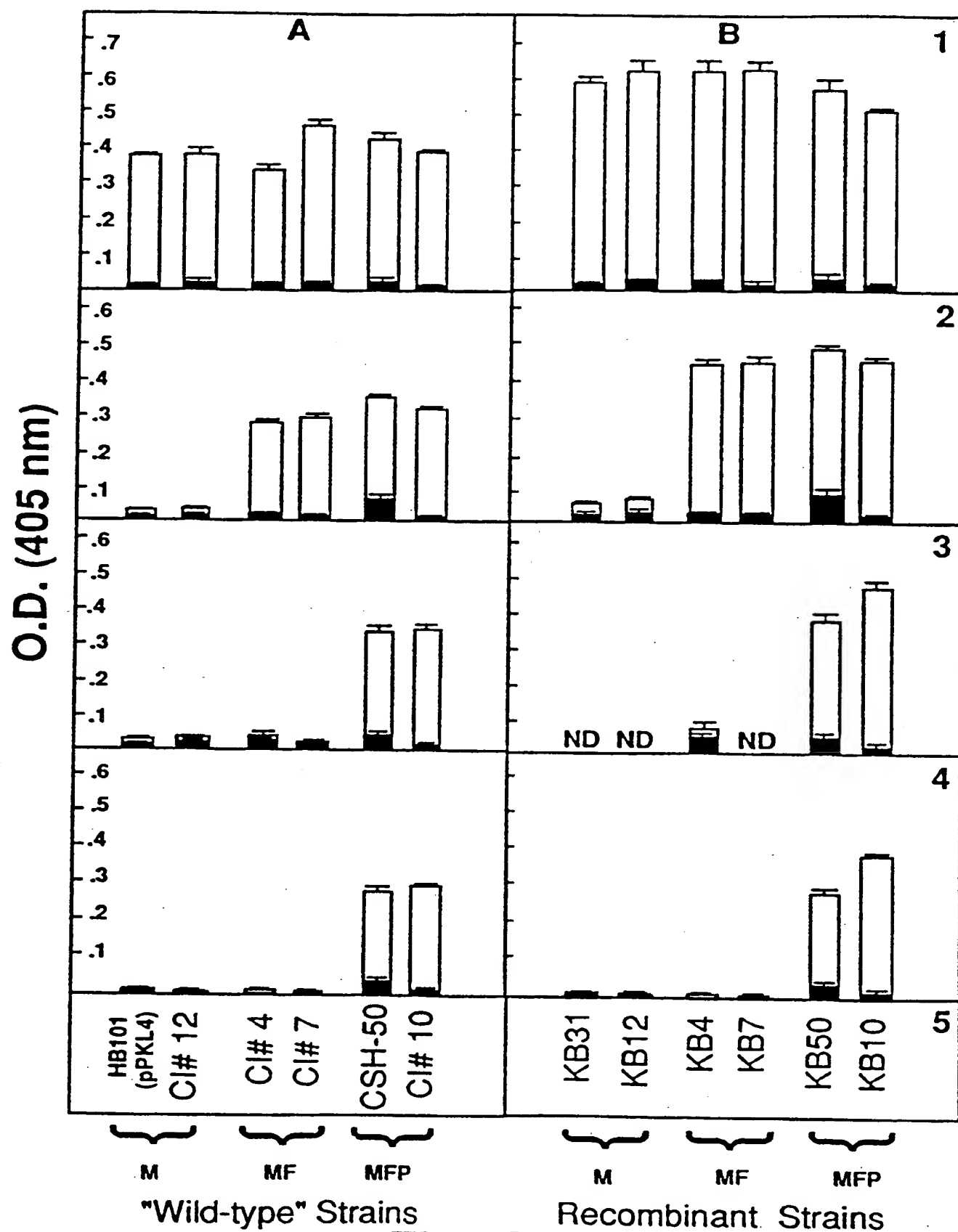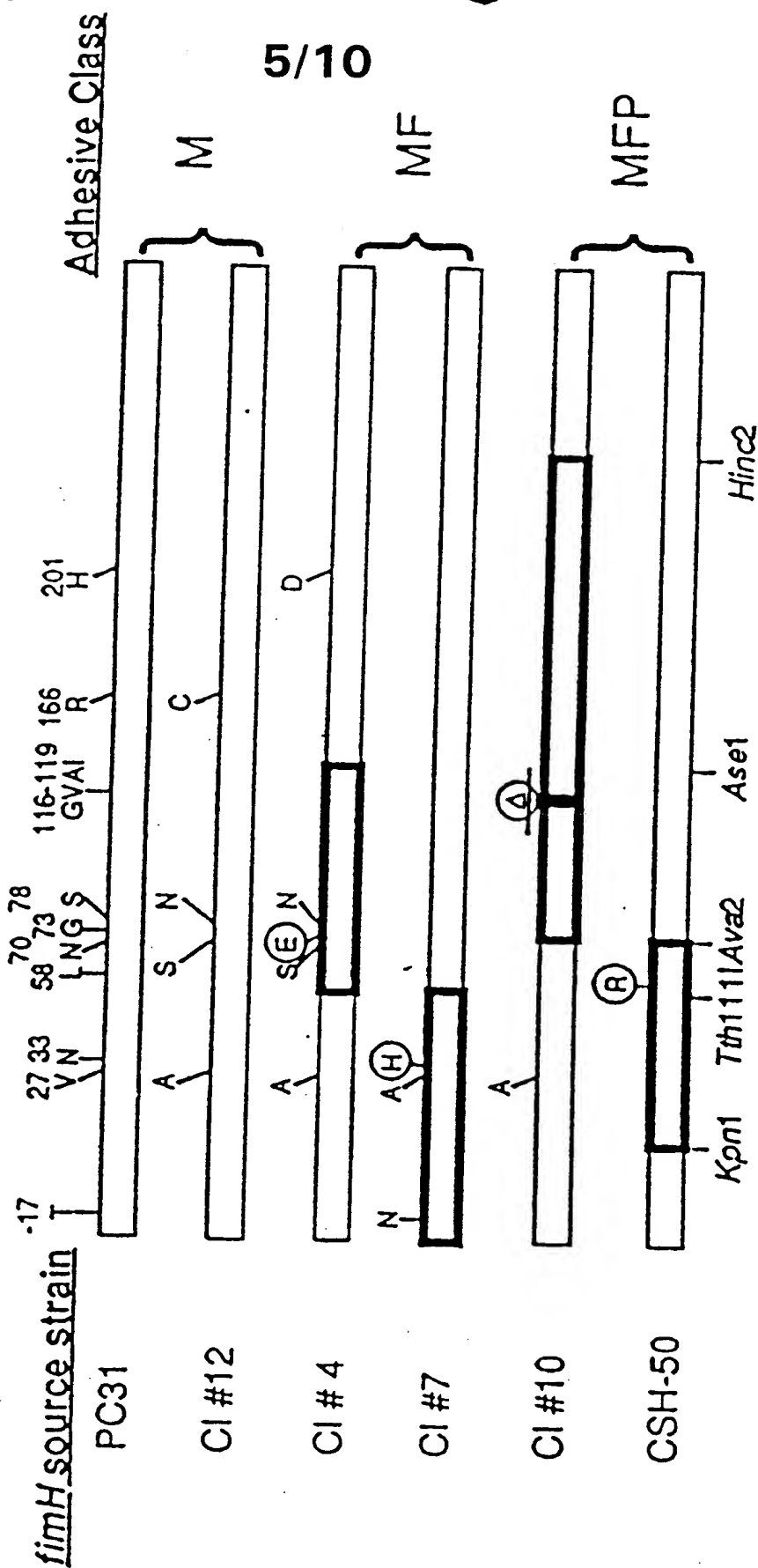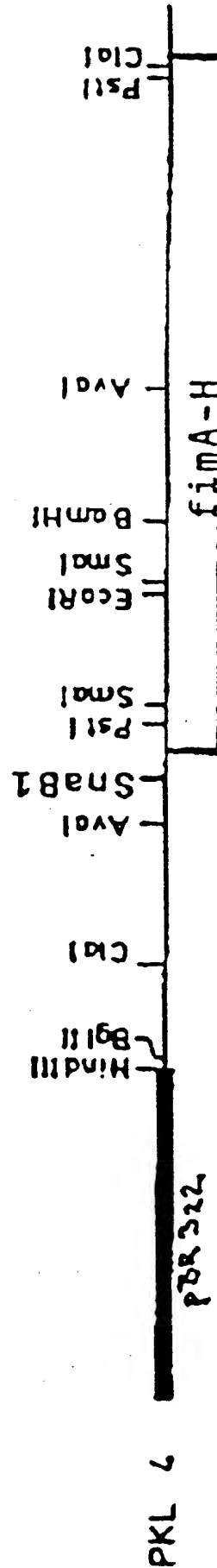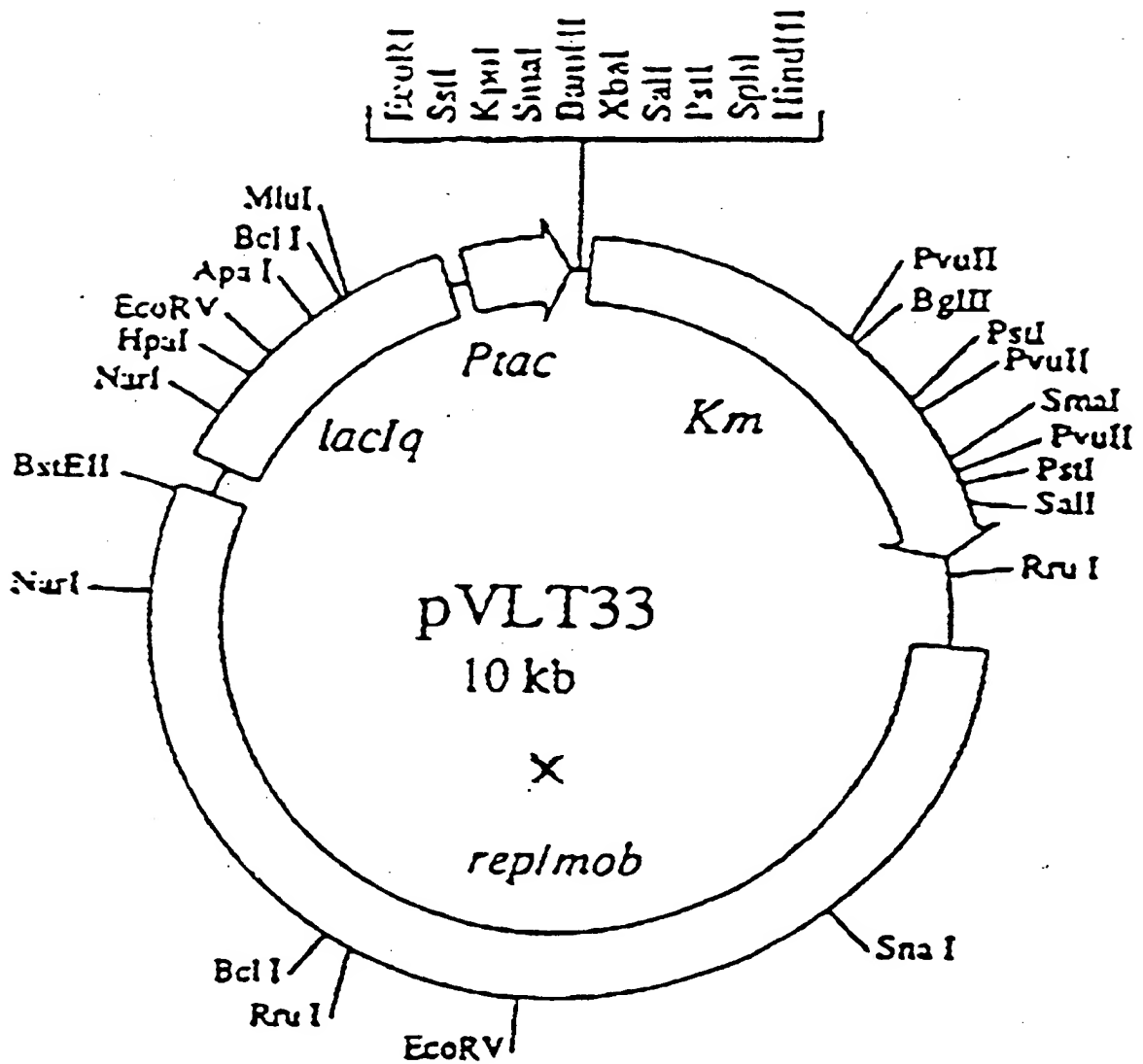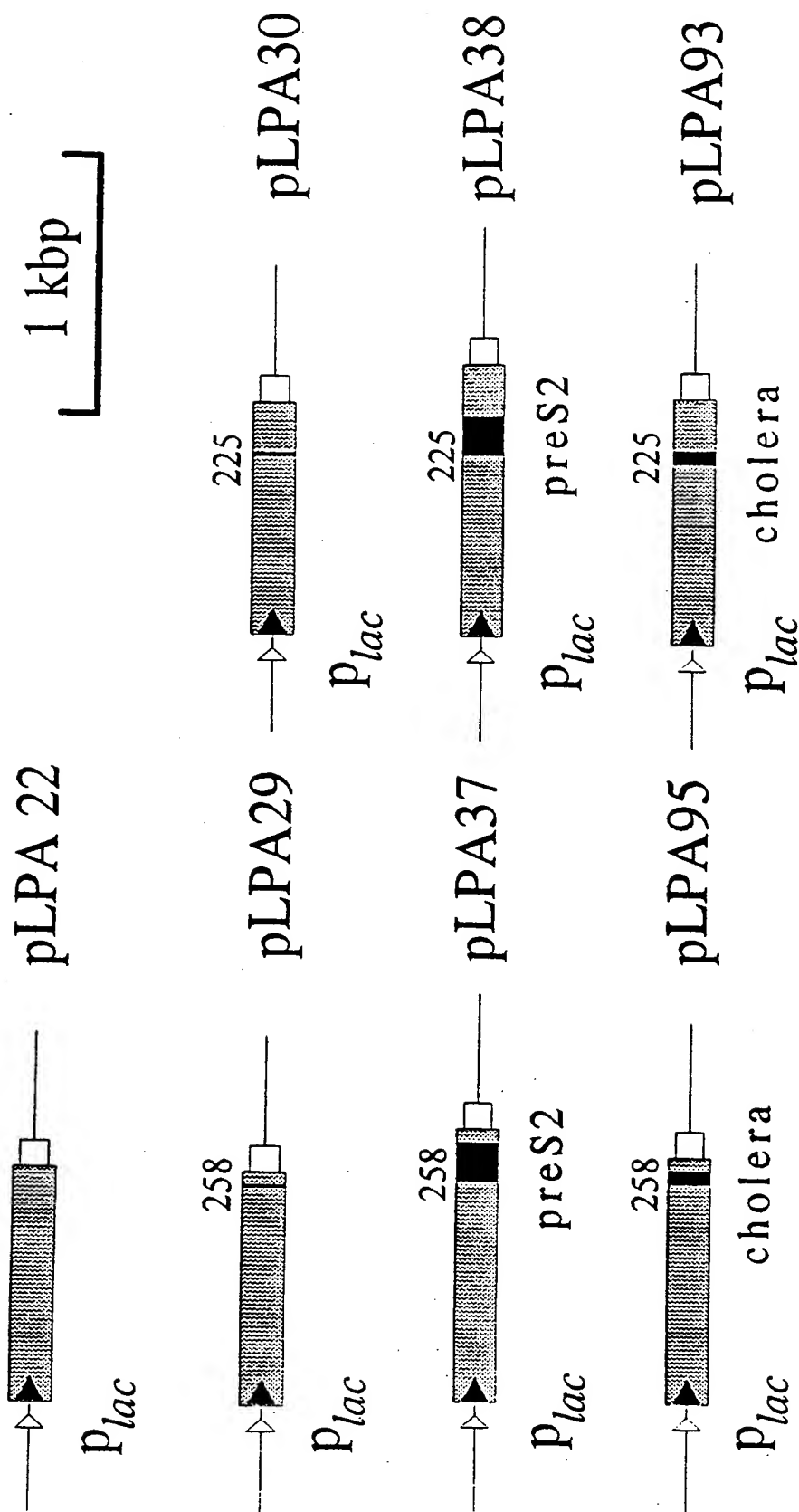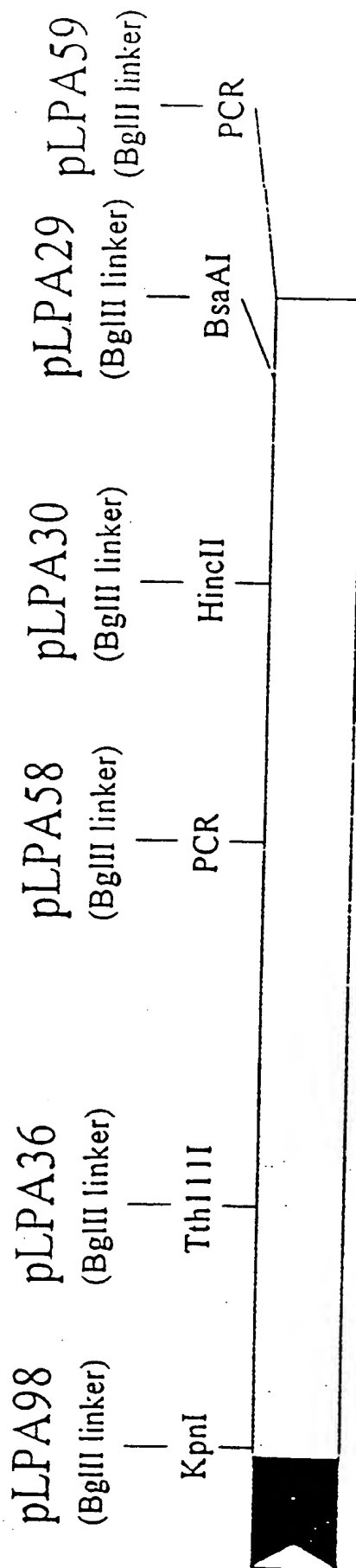
Fig. 1

Fig. 2

Fig. 3

Fig. 4

Fig. 5

Fig. 6

●   ●

Fig. 7

Fig. 8

Fig. 9
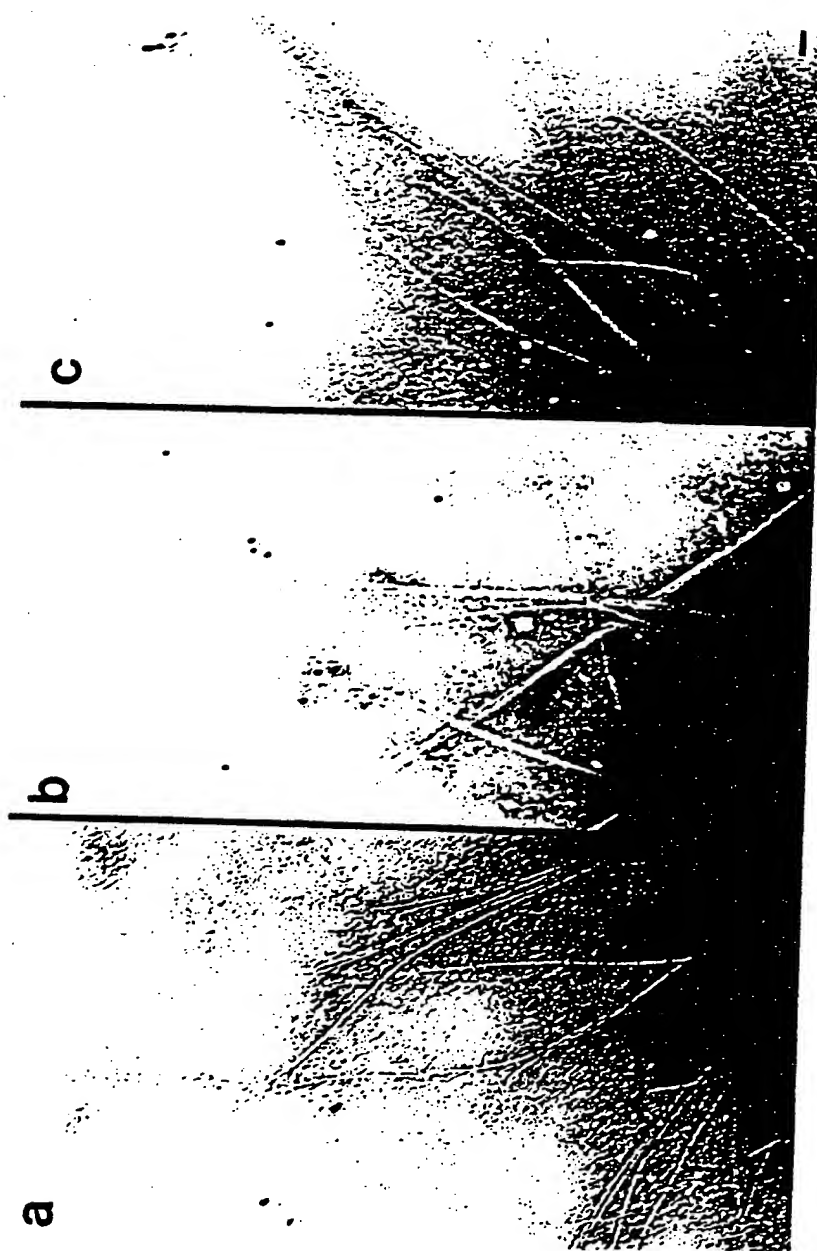
Fig. 10

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6   C12N15/31   C12N15/62   C12N1/21   C07K14/245   C12Q1/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6   C07K   C12N   C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	INFECTION AND IMMUNITY, vol. 60, no. 11, November 1992 WASHINGTON US, pages 4709-4719, EVGENI V. SOKURENKO ET AL. 'Functional heterogeneity of Type 1 fimbriae of Escherichia coli' see abstract see page 4709, right column, last paragraph - page 4710, left column, paragraph 1 see page 4712, right column, paragraph 1 - page 4714, right column, paragraph 1 see page 4715, right column, paragraph 1 - page 4716, left column, paragraph 3 see page 4717, left column, paragraph 2 --- -/--	29,30, 38-40, 43,45, 62,64

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 June 1995	- 3. 07. 95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Montero Lopez, B

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ABSTRACTS OF THE ANNUAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, 16 May 1993 - 20 May 1993 WASHINGTON US, page 113 E. SOKURENKO ET AL. 'Adhesive specificity of Type 1 fimbriae of Escherichia coli. Structural heterogeneity of fimH results in adhesive subclasses' see abstract no. D-101 ---	29-33, 38-40, 43,46,48
X	JOURNAL OF CELLULAR BIOCHEMISTRY. KEYSTONE SYMPOSIA ON MOLECULAR & CELLULAR BIOLOGY. Supplement 17A, January 9-31, 1993; page 376 see abstract no. CZ 301 ---	29-31, 38,50,62
X	WO,A,91 17185 (THE UNITED STATES OF AMERICA) 14 November 1991 see page 3, line 8 - line 16 see page 6, line 4 - line 29 see page 8, line 14 - page 9, line 2 ---	1,2,10
X	WO,A,90 00614 (BAYLOR COLLEGE OF MEDICINE) 25 January 1990 see page 4, line 17 - line 24 see page 5, line 30 - page 6, line 10 see page 6, line 19 - page 7, line 12 see page 18, line 30 - page 20, line 3 see page 20, line 29 - line 34 ---	1,2,4,10
X	EP,A,0 578 293 (AKZO N.V.) 12 January 1994 see page 4, line 31 - line 44 see page 5, line 20 - line 26 see page 8, line 12 - line 19 see page 8, line 45 - line 56 ---	74,77-79
P,X	JOURNAL OF BACTERIOLOGY, vol. 176, no. 3, February 1994 pages 748-755, EVGENI V. SOKURENKO ET AL. 'FimH family of type 1 fimbrial adhesins: Functional heterogeneity due to minor sequence variations among fimH genes' see abstract see page 750, left column, paragraph 2 - page 752, right column, paragraph 3 -----	29-35, 38-52, 54,62-64

1

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9117185	14-11-91	US-A-	5202113	13-04-93
		AU-A-	7858691	27-11-91
WO-A-9000614	25-01-90	NONE		
EP-A-578293	12-01-94	AU-B-	4134393	23-12-93
		CA-A-	2098533	19-12-93
		JP-A-	6056696	01-03-94
		NZ-A-	247907	22-12-94